



**FRIEDRICH-SCHILLER-  
UNIVERSITÄT  
JENA**

# Risikoneutrale Bewertung unter rauer Volatilität

MASTERARBEIT

zur Erlangung des akademischen Grades

Master of Science (M.Sc.)

im Studiengang Wirtschaftsmathematik

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA  
Fakultät für Mathematik und Informatik

eingereicht von Benedikt Heiland

geb. am 22.05.1996 in Erfurt

Betreuer: Prof. Dr. S. Ankirchner

Jena, 30.09.2020

# Abstract

Historische und aktuelle Marktdaten lassen darauf schließen, dass die klassische Annahme einer konstanten Volatilität für die meisten Assets zu stark vereinfachend ist. Stattdessen scheint auch die Volatilität selbst ein stochastischer Prozess zu sein. Im Rahmen dieser Arbeit wird anhand von zwei Datensätzen gezeigt, dass sich der Volatilitätsprozess sehr gut mit Hilfe einer fraktalen Brownschen Bewegung modellieren lässt. Für ein besseres Verständnis wird zunächst die fraktale Brownsche Bewegung formal eingeführt und genauer untersucht. Weiterhin wird gezeigt, wie sich der Volatilitätsprozess eines Assets mit Hilfe der realisierten Volatilität auf Basis von hochfrequenten Marktdaten schätzen lässt. In einem letzten Abschnitt wird untersucht, wie es sich auf die risikoneutrale Bewertung von Derivaten auswirkt, wenn die Volatilität mit Hilfe einer fraktalen Brownschen Bewegung modelliert wird und eine Methode zur schnelleren Kalibrierung des Modells mit Hilfe von künstlichen neuronalen Netzen grob skizziert.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>3</b>
<b>2</b>	<b>Die fraktale Brownsche Bewegung</b>	<b>5</b>
2.1	Definition und Integraldarstellung . . . . .	5
2.2	Elementare Eigenschaften . . . . .	13
2.3	Simulation der Pfade einer fraktalen Brownschen Bewegung . . . . .	16
<b>3</b>	<b>Der Volatilitätsprozess</b>	<b>19</b>
3.1	Stochastische Volatilität . . . . .	19
3.2	Implizite Volatilität und Volatilitätsoberfläche . . . . .	21
3.3	Realisierte Volatilität . . . . .	22
<b>4</b>	<b>Volatilität ist rau</b>	<b>27</b>
4.1	Vorstellung der Datensätze . . . . .	27
4.2	Die Skalierungseigenschaft der Log-Volatilität . . . . .	29
4.2.1	Schätzung des Hurst-Parameters . . . . .	30
4.2.2	Schätzung des Skalierungsfaktors . . . . .	32
4.3	Untersuchung der Inkremente . . . . .	33
<b>5</b>	<b>Risikoneutrale Bewertung unter rauer Volatilität</b>	<b>41</b>
5.1	Die Volatilität unter dem physischen Maß $\mathbb{P}$ . . . . .	41
5.2	Die Volatilität unter dem äquivalenten Martingalmaß $\mathbb{Q}$ . . . . .	44
5.3	Zur Arbitragefreiheit des Marktmodells . . . . .	46
5.4	Simulation des Modells . . . . .	52
5.5	Kalibrierung des Modells . . . . .	56
5.5.1	Anpassung an die Volatilitätsfläche . . . . .	56
5.5.2	Kalibrierung mit Hilfe künstlicher neuronaler Netze . . . . .	57
<b>6</b>	<b>Fazit</b>	<b>62</b>
	<b>Literaturverzeichnis</b>	<b>64</b>

# 1 Einleitung

Im Jahre 1973 stellen Fisher Black und Myron Scholes in [1] das wohl bekannteste Modell zur risikoneutralen Bewertung von Optionen vor. Eine grundlegende Annahme dieses Modells ist eine über die komplette Zeit konstante Volatilität. Wie sich jedoch mit der Hilfe von historischen und aktuellen Marktdaten zeigen lässt, ist diese Annahme zu stark vereinfachend und das Phänomen des Volatility Smile lässt sich damit nicht erklären. Aufgrund dessen wurden bereits mehrere sogenannte stochastische Volatilitätsmodelle vorgestellt, bei denen die Volatilität eines Assets selbst ein stochastischer Prozess ist. Ein weit verbreitetes Modell ist beispielsweise das in [2] präsentierte Heston-Modell. Im Rahmen dieser Arbeit wird ein 2018 von Gatheral, Jaisson und Rosenbaum in [3] vorgestelltes stochastisches Volatilitätsmodell ausführlich präsentiert, welches einige charakteristische Eigenschaften des Volatilitätsprozesses mit nur wenigen Parametern beschreiben kann. Innerhalb dieses Modells wird ein von Benoit Mandelbrot und John Van Ness in [4] eingeführter Prozess, die sogenannte fraktale Brownsche Bewegung, zur Modellierung der stochastischen Volatilität verwendet.

Im ersten Abschnitt der vorliegenden Arbeit wird zunächst die fraktale Brownsche Bewegung formal eingeführt. Weiterhin wird eine Integraldarstellung bezüglich einer Standard Brownschen Bewegung hergeleitet und es werden einige charakteristische Eigenschaften bewiesen. Der darauf folgende Abschnitt befasst sich genauer mit dem Begriff der Volatilität. Es wird gezeigt, weshalb die Annahme von konstanter Volatilität im Widerspruch zu empirischen Beobachtungen steht und damit die Einführung eines stochastischen Volatilitätsprozesses motiviert. Weiterhin werden die Begriffe der impliziten Volatilität sowie der Volatilitätsoberfläche definiert. Bei der empirischen Analyse der historischen Volatilität eines Assets stellt sich die Frage, auf welche Weise diese bestimmt werden soll, da sie sich, im Gegensatz zum Marktpreis, nicht direkt am Markt beobachten lässt. Es wird daher das Konzept der realisierten Volatilität eingeführt, welches historische, hochfrequente Preisdaten verwendet, um den Volatilitätsprozess eines Assets zu schätzen.

Der dritte Abschnitt dieser Arbeit widmet sich der Analyse der realisierten Volatilität von zwei verschiedenen Assets. Es wird die realisierte Volatilität der Microsoft-Aktie und des S&P 500 Index untersucht und es stellt sich heraus, dass diese Prozesse in etwa dieselbe Skalierungseigenschaft wie eine fraktale Brownsche Bewegung aufweisen. Dieser Teil der Arbeit verifiziert hauptsächlich die Ergebnisse, die bereits in [3] vorgestellt wurden. Zusätzlich werden die Inkremente des logarithmierten Vola-

tilitätsprozesses einer genaueren Untersuchung unterzogen und es wird gezeigt, dass die Verteilung ebenjener signifikant von einer Normalverteilung abweicht.

Im vierten und letzten Abschnitt wird untersucht, wie sich Optionen innerhalb eines Modells bewerten lassen, in dem der Volatilitätsprozess mit Hilfe einer fraktalen Brownschen Bewegung modelliert wird. Es wird sich herausstellen, dass das Problem der Pfadabhängigkeit der fraktalen Brownschen Bewegung mit Hilfe von Varianzswaps praktisch gelöst werden kann. Weiterhin wird bewiesen, dass das vorgestellte Marktmodell tatsächlich arbitragefrei ist. Ein Nachteil des betrachteten Modells ist der hohe Rechenaufwand, der nötig ist, um einzelne Realisierungen des Preisprozesses zu simulieren. Deshalb wird, nachdem eine Methode zur Simulation des Modells vorgestellt wurde, zuletzt eine alternative Methode zur schnelleren Kalibrierung des Modells mit Hilfe von künstlichen neuronalen Netzen grob dargestellt werden.

## 2 Die fraktale Brownsche Bewegung

Das Hauptziel dieses ersten Abschnittes ist es, das zentrale Objekt dieser Arbeit einzuführen: die fraktale Brownsche Bewegung. Die fraktale Brownsche Bewegung wurde zunächst von Benoit M. Mandelbrot und John W. Van Ness in [4] vorgestellt und ist eine Verallgemeinerung der klassischen Brownschen Bewegung. Im Gegensatz zur klassischen Brownschen Bewegung sind die Inkremente der fraktalen Brownschen Bewegung im Allgemeinen nicht unabhängig, sondern miteinander korreliert. Auf welche Art und Weise die Inkremente der fraktalen Brownschen Bewegung miteinander korreliert sind hängt dabei von einem Hurst-Parameter  $H \in (0, 1)$  ab. Für  $H < 0.5$  sind die Inkremente der fraktalen Brownschen Bewegung negativ und für  $H > 0.5$  positiv korreliert, wobei mit fallendem (resp. wachsendem)  $H$  die Korrelation abnimmt (resp. zunimmt). Für  $H = 0.5$  entspricht die fraktale Brownsche Bewegung gerade der klassischen Brownschen Bewegung, das heißt die Inkremente sind unkorreliert. Aufgrund dieser Eigenschaft eignet sich die fraktale Brownsche Bewegung, um reale Prozesse zu modellieren, die eine gewisse Pfadabhängigkeit aufweisen. In den folgenden Abschnitten wird die fraktale Brownsche zunächst formal eingeführt und es wird gezeigt, dass diese sich als ein stochastisches Integral bezüglich einer zweiseitigen Brownschen Bewegung darstellen lässt. Anschließend werden einige elementare und für die restliche Arbeit relevante Eigenschaften hergeleitet.

### 2.1 Definition und Integraldarstellung

Da im Folgenden die fraktale Brownsche Bewegung als ein Gaußprozess mit gegebener Kovarianzfunktion eingeführt wird, gilt es sich zunächst noch einmal die Definition eines Gaußprozesses in Erinnerung zu rufen.

**Definition 2.1** (Gaußprozess). *Sei  $(X_t)_{t \in \mathbb{T}}$  ein reellwertiger stochastischer Prozess auf einer beliebigen Indexmenge  $\mathbb{T}$ . Dann heißt  $(X_t)$  Gaußprozess, falls für alle  $n \in \mathbb{N}$  und  $t_1, \dots, t_n \in \mathbb{T}$  der Vektor  $(X_{t_1}, \dots, X_{t_n})$  mehrdimensional normalverteilt ist. Falls  $\mathbb{E}X_t = 0$  für alle  $t \in \mathbb{T}$ , so heißt  $(X_t)$  zentrierter Gaußprozess.*

Im Folgenden wird die fraktale Brownsche Bewegung als ein zentrierter Gaußprozess mit Indexmenge  $\mathbb{T} = \mathbb{R}_{\geq 0}$  und einer speziellen Kovarianzfunktion eingeführt.

**Definition 2.2** (Fraktale Brownsche Bewegung). *Eine fraktale Brownschen Bewegung mit Hurst-Parameter  $H \in (0, 1)$  ist ein zentrierter Gaußprozess  $(B_t^H)_{t \geq 0}$  mit den folgenden Eigenschaften:*

1.  $B_0^H = 0$

$$2. \mathbb{E}[B_t^H B_s^H] = \frac{1}{2} (t^{2H} + s^{2H} - |t - s|^{2H}) =: k(t, s).$$

Die Funktion  $k(t, s)$  wird als Kovarianzfunktion bezeichnet. Eine nützliche Tatsache, welche im weiteren Verlauf der Arbeit noch von Bedeutung sein wird, ist die Integraldarstellung der fraktalen Brownschen Bewegung. Es besteht die Möglichkeit, ein stochastisches Integral bezüglich einer zweiseitigen Standard Brownschen Bewegung zu definieren, welches alle charakteristischen Eigenschaften einer fraktalen Brownschen Bewegung aufweist. Diese Darstellung soll im Folgenden hergeleitet werden. Dazu ist es zunächst notwendig die zweiseitige Brownsche Bewegung sowie stochastische Integrale bezüglich dieser zu definieren.

**Definition 2.3** (Zweiseitige Brownsche Bewegung). *Sei  $(B_t^1, B_t^2)$  eine zweidimensionale Brownsche Bewegung auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ . Dann definiere die zweiseitige Brownsche Bewegung für  $t \in \mathbb{R}$  als  $B_t := B_t^1 1_{t \geq 0} + B_{-t}^2 1_{t < 0}$ . Für deterministische Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f \in L^2((-\infty, t])$  für alle  $t > 0$  definiere*

$$\int_{-\infty}^t f(s) dB_s := \int_0^\infty f(-s) dB_s^2 + \int_0^t f(s) dB_s^1. \quad (1)$$

Die Bedingung  $\int_{-\infty}^t (f(s))^2 ds < \infty$  stellt hierbei sicher, dass die beiden stochastischen Integrale der Summe in Gleichung (1) tatsächlich existieren. Es folgen einige Hilfslemmata, welche benötigt werden um schließlich in Theorem 2.1 zu zeigen, dass das in (1) definierte Integral stets ein Gaußprozess ist.

**Lemma 2.1.** *Sei  $(X_t)_{t \geq 0}$  ein stochastischer Prozess, wobei für alle  $t > 0$  die Zufallsvariable  $X_t$  normalverteilt sei mit  $\mathbb{E}[X_t] = \mu_t$  und  $\text{Var}(X_t) = \sigma_t^2$ . Falls  $\mu_t \rightarrow \mu < \infty$  und  $\sigma_t^2 \rightarrow \sigma^2 < \infty$  für  $t \rightarrow \infty$ , so konvergiert  $X_t$  für  $t \rightarrow \infty$  in Verteilung gegen eine normalverteilte Zufallsvariable  $X_\infty$  mit  $\mathbb{E}[X_\infty] = \mu$  und  $\text{Var}(X_\infty) = \sigma^2$ .*

*Beweis.* Die charakteristische Funktion von  $X_t$  ist gegeben durch

$$\varphi_{X_t}(u) = e^{iu\mu_t - \frac{1}{2}\sigma_t^2 u^2}$$

und damit folgt

$$\varphi_{X_\infty}(u) = e^{iu\mu - \frac{1}{2}\sigma^2 u^2} = \lim_{t \rightarrow \infty} \varphi_{X_t}(u),$$

was gerade die charakteristische Funktion einer normalverteilten Zufallsvariable mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$  ist. Die Aussage folgt jetzt aus der Tatsache, dass

die Konvergenz der charakteristischen Funktionen auch die Konvergenz in Verteilung impliziert.  $\square$

**Lemma 2.2.** Sei  $f \in L^2([0, t])$  für alle  $t > 0$  und  $(B_t)$  eine Standard Brownsche Bewegung, dann ist

$$Y_t := \int_0^t f(s) dB_s$$

ein Gaußprozess.

*Beweis.* Da  $f \in L^2([0, t])$  für alle  $t \geq 0$ , existiert für alle  $t > 0$  eine Folge von Stufenfunktionen  $(f_n)_{n \in \mathbb{N}}$ , so dass

$$\lim_{n \rightarrow \infty} \int_0^t (f(s) - f_n(s))^2 ds = 0$$

und damit insbesondere

$$\lim_{n \rightarrow \infty} \int_0^t f_n(s) dB_s \xrightarrow{L^2} Y_t.$$

Da die Funktionen  $(f_n)$  Stufenfunktionen sind, existieren für jedes  $n \in \mathbb{N}$  ein  $m(n) \in \mathbb{N}$  sowie Konstanten  $c_1^n, \dots, c_{m(n)}^n \in \mathbb{R}$  und  $t_0^n, t_1^n, \dots, t_{m(n)}^n \in [0, t]$  mit  $t_0^n < t_1^n < \dots < t_{m(n)}^n$ , so dass

$$f_n(s) = \sum_{i=1}^{m(n)} c_i^n \mathbb{1}_{(t_{i-1}^n, t_i^n]}(s).$$

Damit gilt dann

$$\int_0^t f_n(s) dB_s = \sum_{i=1}^{m(n)} c_i^n (B_{t_i^n} - B_{t_{i-1}^n})$$

was als Summe unabhängiger, normalverteilter Zufallsvariablen wieder normalverteilt ist. Damit ist also  $\int_0^t f_n(s) dB_s$  normalverteilt mit Erwartungswert 0 und Varianz  $\sum_{i=1}^{m(n)} (c_i^n)^2 (t_i^n - t_{i-1}^n) = \int_0^t (f_n(s))^2 ds$ . Da  $\int_0^t (f_n(s))^2 ds \rightarrow \int_0^t (f(s))^2 ds < \infty$  für  $n \rightarrow \infty$ , ist nach Lemma 2.1 die Zufallsvariable  $Y_t$  ebenfalls normalverteilt.

Es bleibt noch zu zeigen, dass für alle  $0 < t_1 < t_2 < \dots < t_n$  der Vektor  $(Y_{t_1}, \dots, Y_{t_n})^T$



mehrdimensional normalverteilt ist. Dies folgt aus der Tatsache, dass

$$\begin{pmatrix} Y_{t_1} \\ \vdots \\ Y_{t_n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} Y_{t_1} \\ Y_{t_2} - Y_{t_1} \\ \vdots \\ Y_{t_{n-1}} - Y_{t_{n-2}} \\ Y_{t_n} - Y_{t_{n-1}} \end{pmatrix},$$

wobei der Vektor  $(Y_{t_1}, Y_{t_2} - Y_{t_1}, \dots, Y_{t_{n-1}} - Y_{t_{n-2}}, Y_{t_n} - Y_{t_{n-1}})^T$  mehrdimensional normalverteilt ist, da die einzelnen Einträge unabhängig und jeweils normalverteilt sind. Somit ist aber auch  $(Y_{t_1}, \dots, Y_{t_n})^T$  als affine Transformation wieder mehrdimensional normalverteilt.  $\square$

Lemma 2.2 besagt also, dass der zweite Summand in Gleichung (1) ein Gaußprozess ist. Damit das gesamte Integral ein Gaußprozess ist, bleibt noch zu zeigen, dass der erste Summand ebenfalls normalverteilt ist. Wegen der Unabhängigkeit der Summanden folgt dann daraus, dass das stochastische Integral ebenfalls ein Gaußprozess ist.

**Lemma 2.3.** *Sei  $f \in L^2([0, \infty))$  und  $(B_t)$  eine Standard Brownsche Bewegung. Dann ist  $\int_0^\infty f(s)dB_s$  eine normalverteilte Zufallsvariable.*

*Beweis.* Da  $f \in L^2([0, \infty))$ , ist der Prozess  $M_t = \int_0^t f(s)dB_s$  ein Martingal mit  $\mathbb{E}[(M_t)^2] = \int_0^t (f(s))^2 ds \leq \int_0^\infty (f(s))^2 ds < \infty$ . Somit existiert nach dem Martingalkonvergenzsatz eine Zufallsvariable  $M_\infty$ , so dass  $M_t \xrightarrow{L^2} M_\infty$  für  $t \rightarrow \infty$ . Mit Lemma 2.1 folgt, dass  $M_\infty$  normalverteilt ist.  $\square$

Mit diesen drei Lemmata folgt nun, dass das stochastische Integral in Gleichung (1) ein Gaußprozess ist. Für die Integraldarstellung der fraktalen Brownschen Bewegung müssen jedoch stochastische Integrale mit einem allgemeinerem Integranden der Form  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , also der Gestalt  $\int_{-\infty}^t f(t, s)dB_s$ , untersucht werden. Unter welcher Bedingung solche Integrale ebenfalls Gaußprozesse sind, ist Inhalt des folgenden Theorems.

**Theorem 2.1.** *Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  eine deterministische Funktion mit  $f(t, \cdot) \in L^2((-\infty, t])$  für alle  $t \geq 0$  und  $(B_t)$  eine zweiseitige Brownsche Bewegung. Dann ist der Prozess  $(Y_t)$  mit*

$$Y_t = \int_{-\infty}^t f(t, s)dB_s \quad t \geq 0$$

ein Gaußprozess.

*Beweis.* Da  $(B_t)$  eine zweiseitige Brownsche Bewegung ist, gibt es zwei unabhängige Brownsche Bewegungen  $(B_t^1), (B_t^2)$ , so dass

$$Y_t = \int_0^\infty f(t, -s) dB_s^2 + \int_0^t f(t, s) dB_s^1.$$

Der Prozess  $(Y_t)$  ist ein Gaußprozess, falls der Vektor

$$(Y_{t_1}, \dots, Y_{t_n})^T$$

mehrdimensional normalverteilt ist für alle  $n \in \mathbb{N}$  und  $t_1, \dots, t_n \in [0, \infty)$  mit  $t_1 < \dots < t_n$ . Dies ist genau dann der Fall, falls für alle  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  die Linearkombination  $\lambda_1 Y_{t_1}, \dots, \lambda_n Y_{t_n}$  normalverteilt ist. Nun gilt

$$\begin{aligned} & \lambda_1 Y_{t_1} + \dots + \lambda_n Y_{t_n} \\ &= \int_0^\infty \lambda_1 f(t_1, -s) + \dots + \lambda_n f(t_n, -s) dB_s^2 + \int_0^{t_1} \lambda_1 f(t_1, s) + \dots + \lambda_n f(t_n, s) dB_s^1 \\ &+ \int_{t_1}^{t_2} \lambda_2 f(t_2, s) + \dots + \lambda_n f(t_n, s) dB_s^1 + \dots + \int_{t_{n-1}}^{t_n} \lambda_n f(t_n, s) dB_s^1 \\ &=: M_0 + M_1 + \dots + M_n, \end{aligned}$$

wobei  $M_0$  nach Lemma 2.3 eine normalverteilte Zufallsvariable ist, da wegen  $f(t_k, \cdot) \in L^2((-\infty, t_k])$  für alle  $k = 1, \dots, n$  gilt:

$$\begin{aligned} & \int_0^\infty (\lambda_1 f(t_1, -s) + \dots + \lambda_n f(t_n, -s))^2 ds = \int_{-\infty}^0 (\lambda_1 f(t_1, s) + \dots + \lambda_n f(t_n, s))^2 ds \\ & \leq \lambda_1^2 \int_{-\infty}^0 (f(t_1, s))^2 ds + \dots + \lambda_n^2 \int_{-\infty}^0 (f(t_n, s))^2 ds < \infty. \end{aligned}$$

Nach Lemma 2.2 sind die Zufallsvariablen  $M_1, \dots, M_n$  ebenfalls normalverteilt und da  $M_0, M_1, \dots, M_n$  unabhängig sind, folgt dass  $\lambda_1 Y_{t_1} + \dots + \lambda_n Y_{t_n}$  ebenfalls normalverteilt ist.  $\square$

Damit sind die Grundlagen geschaffen, um zu zeigen, dass sich die fraktale Brownsche Bewegung als stochastisches Integral bezüglich einer zweiseitigen Brownschen

Bewegung darstellen lässt.

**Theorem 2.2.** *Sei  $(B_t)$  eine zweiseitige Brownsche Bewegung und  $H \in (0, 1)$ . Dann ist der Prozess*

$$\begin{aligned} B_t^H &= C_H \int_{-\infty}^t (t-s)^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} dB_s \\ &= C_H \left( \int_{-\infty}^0 (t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} dB_s + \int_0^t (t-s)^{H-\frac{1}{2}} dB_s \right) \end{aligned}$$

mit

$$C_H = \left( \frac{2H}{\left(\frac{3}{2} - H\right) B(2-2H, H + \frac{1}{2})} \right)^{\frac{1}{2}}$$

eine fraktale Brownsche Bewegung.

Bei der Funktion  $B(x, y)$  handelt es sich um die Eulersche Betafunktion welche für  $x, y > 0$  wie folgt definiert ist:

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt.$$

*Beweis.* Der Beweis folgt den Darstellungen aus [5, Theorem B.1] und [6, Kapitel 1.2]. Zunächst ist zu zeigen, dass es sich bei dem gegebenen Prozess tatsächlich um einen Gaußprozess handelt. Nach Theorem 2.1 ist dies der Fall, falls  $f(t, \cdot) \in L^2((-\infty, t])$  für alle  $t \geq 0$ , wobei hier  $f(t, s) := (t-s)^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}}$ . Es gilt

$$\begin{aligned} & \int_{-\infty}^t (f(t, s))^2 ds \\ &= \int_{-\infty}^0 \left( (t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right)^2 ds + \int_0^t \left( (t-s)^{H-\frac{1}{2}} \right)^2 ds \\ &= \int_0^{\infty} \left( (t+s)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}} \right)^2 ds + \frac{1}{2H} t^{2H} \\ &= : \Phi(t) + \frac{1}{2H} t^{2H}. \end{aligned}$$

Mit dem Mittelwertsatz der Differentialrechnung folgt für alle  $s, t \geq 0$

$$(t+s)^{H-1/2} - s^{H-1/2} \leq t \cdot \sup_{u \in [s, s+t]} \frac{du^{H-1/2}}{du} = t(H-1/2)s^{H-3/2}$$

und damit gilt

$$\begin{aligned} \Phi(t) &= \int_0^1 ((t+s)^{H-1/2} - s^{H-1/2})^2 ds + \int_1^\infty ((t+s)^{H-1/2} - s^{H-1/2})^2 ds \\ &\leq \int_0^1 (t+s)^{2H-1} ds + \int_1^\infty t^2 (H-1/2)^2 s^{2H-3} ds \\ &= \frac{1}{2H} ((t+1)^{2H} - t^{2H}) + t^2 (H-1/2)^2 \frac{1}{2-2H} \\ &< \infty. \end{aligned}$$

Somit ist  $f \in L^2((-\infty, t])$  für alle  $t \geq 0$  und es bleibt zu zeigen, dass das stochastische Integral tatsächlich die Kovarianzfunktion einer fraktalen Brownschen Bewegung aufweist. Für die erste und zweite Ableitung von  $\Phi(t)$  folgt mit der Leibniz-Regel für Parameterintegrale:

$$\begin{aligned} \Phi'(t) &= \int_0^\infty 2 \left( (t+s)^{H-\frac{1}{2}} - s^{H-\frac{1}{2}} \right) \left( H - \frac{1}{2} \right) (t+s)^{H-\frac{3}{2}} ds \\ &= (2H-1) \int_0^\infty (t+s)^{2H-2} - s^{H-\frac{1}{2}} (t+s)^{H-\frac{3}{2}} ds \\ \Phi''(t) &= (2H-1) \int_0^\infty (2H-2)(t+s)^{2H-3} - s^{H-\frac{1}{2}} \left( H - \frac{3}{2} \right) (t+s)^{H-\frac{5}{2}} ds \\ &= -(2H-1)t^{2H-2} - (2H-1) \left( H - \frac{3}{2} \right) \int_0^\infty s^{H-\frac{1}{2}} (t+s)^{H-\frac{5}{2}} ds. \end{aligned}$$

Nun gilt

$$\begin{aligned} &\int_0^\infty s^{H-\frac{1}{2}} (t+s)^{H-\frac{5}{2}} ds \\ &= t^{2H-3} \int_0^\infty \left( \frac{s}{t} \right)^{H-\frac{1}{2}} \left( 1 + \frac{s}{t} \right)^{H-\frac{5}{2}} ds \end{aligned}$$

$$\begin{aligned}
&= t^{2H-2} \int_1^\infty (u-1)^{H-\frac{1}{2}} u^{H-\frac{5}{2}} du \\
&= t^{2H-2} \int_0^1 \left( \frac{1-z}{z} \right)^{H-\frac{1}{2}} \left( \frac{1}{z} \right)^{H-\frac{5}{2}} \frac{1}{z^2} dz \\
&= t^{2H-2} \int_0^1 (1-z)^{H-\frac{1}{2}} z^{1-2H} dz \\
&= t^{2H-2} B \left( 2-2H, H+\frac{1}{2} \right),
\end{aligned}$$

wobei  $u = 1 + \frac{s}{t}$  und  $z = \frac{1}{u}$  substituiert wurde. Es gilt somit

$$\Phi''(t) = - (2H-1)t^{2H-2} + (2H-1) \left( \frac{3}{2} - H \right) B \left( 2-2H, H+\frac{1}{2} \right) t^{2H-2}$$

und da  $\Phi(0) = \Phi'(0) = 0$  ergibt sich durch zweimaliges Integrieren

$$\Phi'(t) = -t^{2H-1} + \left( \frac{3}{2} - H \right) B \left( 2-2H, H+\frac{1}{2} \right) t^{2H-1}$$

sowie

$$\Phi(t) = -\frac{1}{2H}t^{2H} + \frac{\left( \frac{3}{2} - H \right) B \left( 2-2H, H+\frac{1}{2} \right)}{2H} t^{2H}.$$

Es folgt also

$$\int_{-\infty}^t (f(t,s))^2 ds = \Phi(t) + \frac{1}{2H}t^{2H} = \frac{\left( \frac{3}{2} - H \right) B \left( 2-2H, H+\frac{1}{2} \right)}{2H} t^{2H} = \frac{1}{(C_H)^2} t^{2H}.$$

und insbesondere gilt

$$\mathbb{E} \left[ (B_t^H)^2 \right] = t^{2H}.$$

Es gilt für alle  $t, s > 0$  wegen der zweiten binomischen Formel

$$\mathbb{E} [B_t^H B_s^H] = \frac{1}{2} \left( \mathbb{E} (B_t^H)^2 + \mathbb{E} (B_s^H)^2 - \mathbb{E} (B_t^H - B_s^H)^2 \right)$$

und für  $t > s > 0$  gilt

$$\begin{aligned} B_t^H - B_s^H &= C_H \left( \int_{-\infty}^t (t-u)^{H-\frac{1}{2}} - (-u)_+ dB_u - \int_{-\infty}^s (s-u)^{H-\frac{1}{2}} - (-u)_+ dB_u \right) \\ &= C_H \left( \int_{-\infty}^s (t-u)^{H-\frac{1}{2}} - (s-u)^{H-\frac{1}{2}} dB_u + \int_s^t (t-u)^{H-\frac{1}{2}} dB_u \right). \end{aligned}$$

Somit folgt

$$\begin{aligned} &\mathbb{E} (B_t^H - B_s^H)^2 \\ &= C_H^2 \left( \int_{-\infty}^s \left( (t-u)^{H-\frac{1}{2}} - (s-u)^{H-\frac{1}{2}} \right)^2 du + \int_s^t \left( (t-u)^{H-\frac{1}{2}} \right)^2 du \right) \\ &= C_H^2 \left( \int_{-\infty}^0 \left( (t-s-z)^{H-\frac{1}{2}} - (-z)^{H-\frac{1}{2}} \right)^2 dz + \int_0^{t-s} \left( (t-s-z)^{H-\frac{1}{2}} \right)^2 dz \right) \\ &= C_H^2 \int_{-\infty}^{t-s} \left( (t-s-z)^{H-\frac{1}{2}} - (-z)_+^{H-\frac{1}{2}} \right)^2 dz \\ &= \mathbb{E} (B_{t-s}^H)^2 \\ &= (t-s)^{2H}, \end{aligned}$$

wobei  $z = u - s$  substituiert wurde, und damit gilt

$$\mathbb{E} [B_t^H B_s^H] = \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H}).$$

□

## 2.2 Elementare Eigenschaften

Nachdem die fraktale Brownsche Bewegung definiert und eine Integraldarstellung hergeleitet wurde, sollen in diesem Unterabschnitt einige elementare Eigenschaften präsentiert werden, die im weiteren Verlauf der Arbeit benötigt werden. Zunächst wird gezeigt, dass die Inkremente der fraktalen Brownschen Bewegung stationär sind, das heißt die Verteilung der Inkremente ist invariant gegenüber einer Verschiebung der Indizes.

**Theorem 2.3** (stationäre Zuwächse). *Die fraktale Brownsche Bewegung  $(B_t^H)$  besitzt stationäre Zuwächse, das heißt für alle  $s > 0$  besitzen die beiden Prozesse*

$(B_{t+s}^H - B_s^H)$  und  $(B_t^H)$  dieselbe endlichdimensionale Verteilung.

*Beweis.* Da auch der Prozess  $(B_{t+s}^H - B_s^H)$  ein zentrierter Gaußprozess ist, genügt es zu zeigen, dass er die selbe Kovarianzfunktion wie  $(B_t^H)$  besitzt. Nun gilt für alle  $t_1, t_2 > 0$ :

$$\begin{aligned} & \mathbb{E}[(B_{t_1+s}^H - B_s^H)(B_{t_2+s}^H - B_s^H)] \\ &= \mathbb{E}[B_{t_1+s}^H B_{t_2+s}^H] - \mathbb{E}[B_{t_1+s}^H B_s^H] - \mathbb{E}[B_{t_2+s}^H B_s^H] + \mathbb{E}[B_s^H B_s^H] \\ &= \frac{1}{2}(t_1^{2H} + t_2^{2H} - |t_2 - t_1|^{2H}) \\ &= \mathbb{E}[B_{t_1}^H B_{t_2}^H]. \end{aligned}$$

□

Eine charakteristische Eigenschaft der fraktalen Brownschen Bewegung ist ihre Selbstähnlichkeit. Das bedeutet, dass sich der Prozess so skalieren lässt, dass der skalierte Prozesse wieder dieselbe Verteilung besitzt. Im Falle der fraktalen Brownschen Bewegung hängt der Skalierungsparameter vom Hurst-Parameter  $H$  ab, wie das folgende Theorem zeigt.

**Theorem 2.4** (Selbstähnlichkeit). *Die fraktale Brownsche Bewegung  $(B_t^H)$  ist selbstähnlich mit Hurst-Index  $H$ , das heißt für alle  $c > 0$  besitzen die Prozesse  $(B_{ct}^H)$  und  $c^H B_t^H$  dieselbe endlichdimensionale Verteilung.*

*Beweis.* Da sowohl  $(B_{ct}^H)$  als auch  $(c^H B_t^H)$  ein zentrierter Gaußprozess ist, genügt es zu zeigen, dass beide Prozesse dieselbe Kovarianzfunktion besitzen. Nun gilt aber für alle  $s, t > 0$ :

$$\begin{aligned} \mathbb{E}[B_{ct}^H B_{cs}^H] &= \frac{1}{2}((ct)^{2H} + (cs)^{2H} - |ct - cs|^{2H}) \\ &= \frac{c^{2H}}{2}(t^{2H} + s^{2H} - |t - s|^{2H}) \\ &= \mathbb{E}[c^H B_t^H c^H B_s^H]. \end{aligned}$$

□

Im weiteren Verlauf dieser Arbeit wird stets vorausgesetzt, dass die fraktale Brownsche Bewegung stetige Pfade besitzt, weshalb zu zeigen ist, dass stets eine entsprechende Modifikation existiert. Dazu wird das folgende Theorem von Kolmogorov-Tschenow verwendet.

**Theorem 2.5** (Kolmogorov-Tschenow). *Sei  $(X_t)_{t \geq 0}$  ein reellwertiger stochastischer Prozess. Angenommen für jedes  $T > 0$  existieren reelle Zahlen  $\alpha, \beta, C > 0$ , so dass*

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C|t - s|^{1+\beta} \quad \forall s, t \in [0, T].$$

*Dann besitzt  $(X_t)$  eine Modifikation mit lokal  $\gamma$ -hölderstetigen Pfaden für alle  $0 < \gamma < \frac{\beta}{\alpha}$ .*

*Beweis.* Siehe [7, Theorem 21.6]. □

Mit Hilfe dieses Theorems kann nun gezeigt werden, dass eine fraktale Brownsche Bewegung stets eine Modifikation mit stetigen Pfaden besitzt.

**Theorem 2.6.** *Sei  $(B_t^H)$  eine fraktale Brownsche Bewegung. Dann besitzt  $(B_t^H)$  eine Modifikation  $(\tilde{B}_t^H)$  mit lokal  $\gamma$ -hölderstetigen Pfaden für alle  $0 < \gamma < H$ .*

*Beweis.* Der Beweis folgt der Darstellung von [6, Theorem 1.6.1]. Nach dem Theorem von Kolmogorov-Tschenow besitzt  $(B_t^H)$  eine lokal hölderstetige Modifikation, falls es positive Konstanten  $\alpha, \beta, K$  gibt, so dass für alle  $s, t > 0$  gilt

$$\mathbb{E}[(B_t^H - B_s^H)^\alpha] \leq K|t - s|^{1+\beta}.$$

Nun gilt für alle  $s, t, \alpha > 0$  wegen der Stationarität der Inkremente und der Selbstähnlichkeitseigenschaft

$$\mathbb{E}[(B_t^H - B_s^H)^\alpha] = \mathbb{E}[|B_{t-s}^H|^\alpha] = \mathbb{E}[|B_1^H|^\alpha] |t - s|^{\alpha H}$$

und somit besitzt  $(B_t^H)$  eine Modifikation mit stetigen und insbesondere lokal  $\gamma$ -hölderstetigen Pfaden für

$$0 < \gamma < \frac{\alpha H - 1}{\alpha} = H - \frac{1}{\alpha}.$$

Da  $\alpha > 0$  beliebig ist, sind die Pfade der Modifikation lokal  $\gamma$ -hölderstetig für alle  $0 < \gamma < H$ . □

Da eine fraktale Brownsche Bewegung somit stets eine Modifikation mit lokal hölderstetigen Pfaden besitzt, ist mit der fraktalen Brownschen Bewegung ab sofort die lokal hölderstetige Modifikation gemeint. Bereits in der Einleitung wurde erwähnt, dass die Inkremente der fraktalen Brownschen Bewegung für  $H < 0.5$  stets negativ und für  $H > 0.5$  stets positiv korreliert sind. Im ersten Fall wird die fraktale Brownsche Bewegung auch als antipersistent und im zweiten Fall als persistent bezeichnet.



**Lemma 2.4.** Sei  $(B_t^H)$  eine fraktale Brownsche Bewegung. Dann gilt für  $0 < s_1 < t_1 < s_2 < t_2$ :

$$\mathbb{E} [(B_{t_2}^H - B_{s_2}^H)(B_{t_1}^H - B_{s_1}^H)] \begin{cases} > 0 \text{ für } H > 0.5 \\ < 0 \text{ für } H < 0.5 \end{cases}$$

*Beweis.* Der Beweis folgt in etwa der Darstellung in [8, S. 2]. Mit der Kovarianzfunktion der fraktalen Brownschen Bewegung folgt

$$\begin{aligned} & \mathbb{E} [(B_{t_2}^H - B_{s_2}^H)(B_{t_1}^H - B_{s_1}^H)] \\ &= \mathbb{E} [B_{t_2}^H B_{t_1}^H] - \mathbb{E} [B_{t_2}^H B_{s_1}^H] - \mathbb{E} [B_{s_2}^H B_{t_1}^H] + \mathbb{E} [B_{s_2}^H B_{s_1}^H] \\ &= \frac{1}{2} ((t_2 - s_1)^{2H} - (t_2 - t_1)^{2H} - ((s_2 - s_1)^{2H} - (s_2 - t_1)^{2H})) \\ &= \frac{1}{2} (f(b_2) - f(a_2) - (f(b_1) - f(a_1))), \end{aligned}$$

wobei  $f(x) = x^{2H}$ . Nun gilt  $b_2 - a_2 = b_1 - a_1 = t_1 - s_1$  sowie  $b_2 > b_1$  und damit folgt die Behauptung, da  $f$  für  $H > 0.5$  konvex und für  $H < 0.5$  konkav ist.  $\square$

## 2.3 Simulation der Pfade einer fraktalen Brownschen Bewegung

In diesem Abschnitt soll eine einfache Methode vorgestellt werden, mit der sich die Pfade einer fraktalen Brownschen Bewegung simulieren lassen. Die Idee des behandelten Verfahrens ist es, die fraktale Brownsche Bewegung an ausgewählten Stützstellen als einen Gaußschen Vektor zu betrachten. Das Verfahren hat jedoch den Nachteil, dass es äußerst rechenintensiv ist. Für einen effizienteren Algorithmus, welcher die Stationarität der Inkremente ausnutzt, sei auf [9] verwiesen.

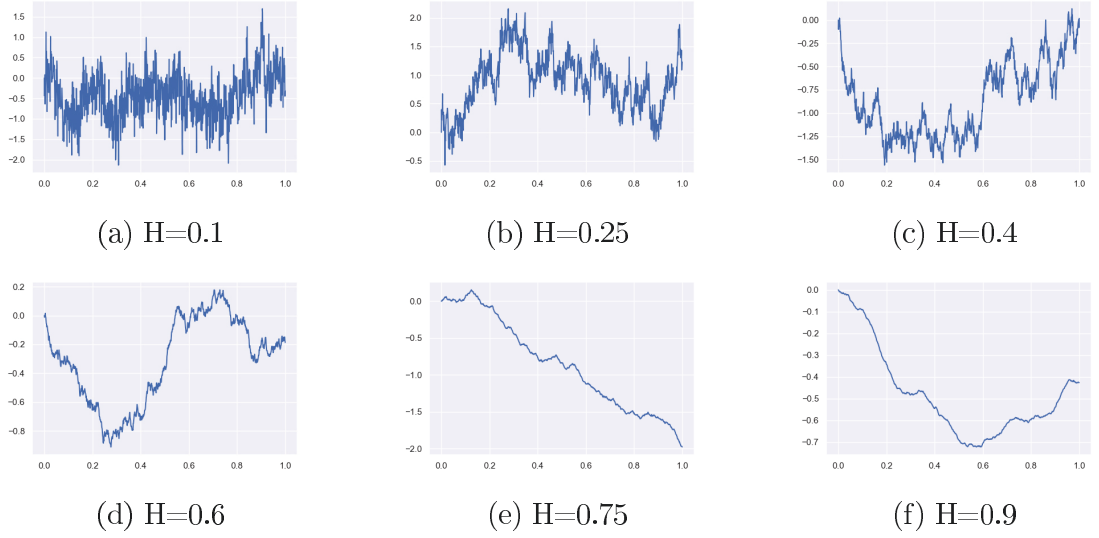


Abbildung 1: Simulierte Pfade einer fraktalen Brownschen Bewegung für verschiedene Hurst-Parameter ( $T=1$ ,  $N=1000$ )

Sei also  $T > 0$  und  $N \in \mathbb{N}$  die Zahl der Stützstellen, an denen  $(B_t^H)_{t \in [0, T]}$  simuliert werden soll. Das Ziel ist eine Realisierung des Vektors  $\left(B_{\frac{T}{N}}^H, B_{\frac{2T}{N}}^H, \dots, B_{\frac{NT}{N}}^H\right)^T$  zu erzeugen. Dazu kann ausgenutzt werden, dass wegen der in Theorem 2.4 gezeigten Selbstähnlichkeitseigenschaft gilt:

$$B_{\frac{kT}{N}}^H \stackrel{d}{=} \left(\frac{T}{N}\right)^{2H} B_k^H.$$

Es genügt also den Gaußschen Vektor  $(B_1^H, B_2^H, \dots, B_N^H)^T$  zu simulieren und anschließend entsprechend zu skalieren. Dazu ist es zunächst erforderlich die Kovarianzmatrix  $\Sigma$  zu berechnen mit  $\Sigma_{mn} = m^{2H} + n^{2H} - |m - n|^{2H}$  für  $m, n \in \{1, \dots, N\}$ . Da eine Kovarianzmatrix stets positiv semidefinit ist, lässt sich mit Hilfe der Cholesky-Zerlegung eine Matrix  $G$  bestimmen, so dass  $GG^T = \Sigma$ . Falls nun  $Z = (Z_1, \dots, Z_N)^T$  ein Vektor von unabhängigen, standard normalverteilten Zufallsvariablen ist, so ist  $GZ$  ein gaußscher Vektor mit Kovarianzmatrix  $\Sigma$  (vgl. [10, S. 307]). Das gesamte Verfahren lässt sich wie folgt zusammenfassen:

1. Berechne die Kovarianzmatrix  $\Sigma$  mit  $\Sigma_{mn} = m^{2H} + n^{2H} - |m - n|^{2H}$  für alle  $m, n \in \{1, \dots, N\}$  und bestimme die Cholesky-Zerlegung  $G$  von  $\Sigma$ .
2. Erzeuge einen Vektor  $Z = (Z_1, \dots, Z_N)$  von unabhängigen standard normalverteilten Zufallsvariablen und berechne  $Y = GZ$ .
3. Berechne  $B^H = \left(\frac{T}{N}\right)^{2H} Y$ .

In Abbildung 1 sind die simulierten Pfade einer fraktalen Brownschen zu sehen. Es ist sehr gut zu erkennen, dass für  $H < 0.5$  die Pfade ein antipersistentes Verhalten aufweisen, das heißt auf einen Anstieg folgt typischerweise ein Abschwung und umgekehrt. Dies führt dazu, dass die Pfade der fraktalen Brownschen Bewegung rau erscheinen und stark hin und her schwanken. Für  $H > 0.5$  sind die Pfade hingegen eher glatt und das persistente Verhalten ist gut zu erkennen. Ein grundlegendes Ziel dieser Arbeit ist, zu zeigen, dass der Volatilitätsprozess von einigen Assets sehr gut mit Hilfe einer fraktalen Brownschen Bewegung dargestellt werden kann. Im folgenden Abschnitt werden daher zunächst einige grundlegende Begrifflichkeiten rund um den Volatilitätsprozess eines Assets ausführlich besprochen.

### 3 Der Volatilitätsprozess

In Abschnitt 4 dieser Arbeit wird der Volatilitätsprozess von zwei Assets untersucht. Deshalb werden in diesem Abschnitt zunächst einige grundlegende und für den weiteren Verlauf relevante Konzepte vorgestellt. Zunächst wird dargestellt, weshalb es sinnvoll ist, die Annahme einer konstanten Volatilität zugunsten einer stochastischen Volatilität aufzugeben. Weiterhin wird gezeigt, wie sich der Volatilitätsprozess mit Hilfe von hochfrequenten Daten schätzen lässt, da dieser Prozess selbst nicht am Markt beobachtbar ist. Diese Schätzung wird auch als realisierte Volatilität bezeichnet. Weiterhin werden die Begriffe der impliziten Volatilität und der Volatilitätsoberfläche genauer erläutert.

#### 3.1 Stochastische Volatilität

Das wohl bekannteste Modell zur Bewertung von Optionen ist das 1973 von Fisher Black und Myron Scholes in [1] vorgestellte Modell, in welchem der Preisprozess  $(S_t)$  eines Assets durch die stochastische Differentialgleichung

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

modelliert wird, wobei  $\mu \in \mathbb{R}$  und  $\sigma > 0$ . Der Parameter  $\sigma$  wird auch als Volatilität bezeichnet und beschreibt wie groß die Schwankungsbreite des Preisprozesses ist. Innerhalb des Black-Scholes-Modells wird angenommen, dass die Volatilität  $\sigma$  des Preisprozesses im Verlauf der Zeit konstant ist und sich somit nicht ändert. Dies impliziert insbesondere, dass die Log-Renditen des Assets normalverteilt sind, denn mit der Ito-Formel für Ito-Prozesse folgt:

$$\log(S_{t+\Delta}/S_t) = \left(\mu - \frac{\sigma^2}{2}\right) \Delta + \sigma(B_{t+\Delta} - B_t) \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right) \Delta, (\Delta\sigma)^2\right).$$

Bereits 1963 zeigte Mandelbrot in [11], dass diese Annahme nicht von den vorhandenen Daten gestützt wird und auch Black und Scholes schrieben in [12]: „[...] *there is evidence of non-stationarity in the variance. More work must be done to predict variances using the information available.*“ Ein Blick auf das Histogramm (Abb. 3) der täglichen Log-Renditen (Eröffnungs- bis Schlusskurs) des S&P 500 Index macht deutlich, dass diese im Vergleich zu einer Normalverteilung beträchtlich spitzer verteilt sind. Zur Veranschaulichung ist in der Abbildung ebenfalls die Dichte einer Normalverteilung zu sehen, wobei Erwartungswert und Varianz aus den gegebenen Daten geschätzt wurden. Es ist zu erkennen, dass sehr große und sehr niedrige Log-

Renditen im Vergleich zu einer Normalverteilung deutlich häufiger auftreten. Der zeitliche Verlauf der Log-Renditen, welcher in Abbildung 2 zu sehen ist, lässt ebenfalls vermuten, dass die Volatilität nicht konstant, sondern von sich abwechselnden Phasen niedriger und hoher Volatilität geprägt ist. Dies lässt sich daran erkennen, dass betragsmäßig eher hohe bzw. niedrige Renditen phasenweise und nicht vollkommen willkürlich auftreten, das heißt auf eine betragsmäßig hohe (resp. niedrige) Rendite folgt mit hoher Wahrscheinlichkeit wieder eine betragsmäßig hohe (resp. niedrige) Rendite (vgl. [11]). Dieses Phänomen wird auch als Volatilitätsclustering bezeichnet.

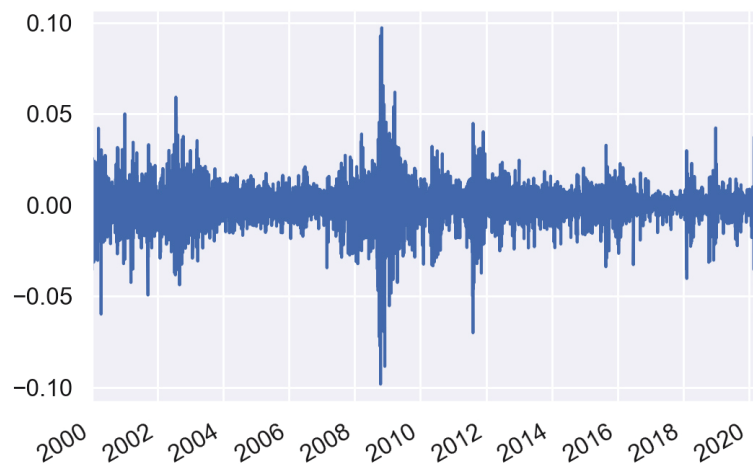


Abbildung 2: Tägliche Log-Renditen des S&P 500 Index im zeitlichen Verlauf

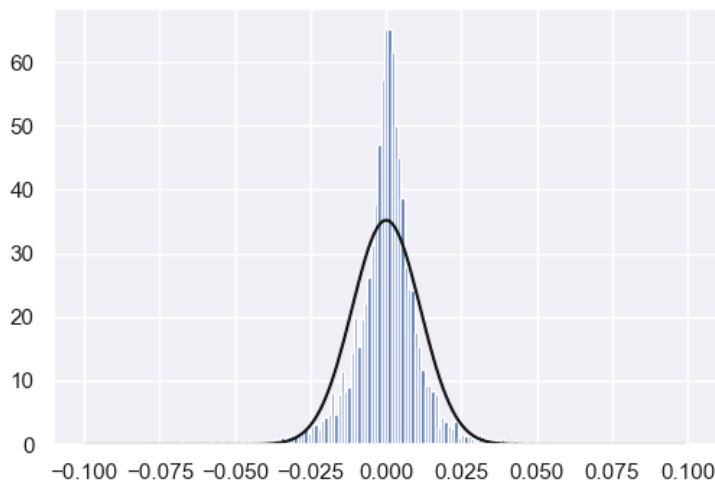


Abbildung 3: Histogramm der täglichen Log-Renditen des S&P 500 Index

Aufgrund dessen lässt sich darauf schließen, dass die Volatilität eines Assets nicht, wie im Black-Scholes-Modell angenommen, eine konstante Größe ist, sondern sich im Zeitverlauf ändert. Es erscheint also sinnvoll, die stochastische Differentialgleichung für den Preisprozess wie folgt abzuändern:

$$dS_t = \mu S_t dt + \sigma_t S_t dB_t.$$

Hierbei ist  $\sigma_t$  nun selbst ein stochastischer Prozess. Gelegentlich wird auch angenommen, dass  $\sigma_t = \sigma(t, S_t)$  eine deterministische Funktion ist, die lediglich von der Zeit und dem aktuellen Preis abhängt. In diesem Fall wird auch von einem lokalen Volatilitätsmodell gesprochen. Der große Vorteil von stochastischen Volatilitätsmodellen liegt darin, dass sich mit ihnen sowohl die Verteilung der Log-Renditen als auch das Phänomen des Volatilitätsclusterings beschreiben lassen. Ein spezielles stochastisches Volatilitätsmodell, welches viele empirisch beobachtbare Eigenschaften des Volatilitätsprozesses mit nur wenigen Parametern beschreiben kann, wird im Rahmen dieser Arbeit im Abschnitt 5 vorgestellt werden.

### 3.2 Implizite Volatilität und Volatilitätsoberfläche

Nachdem erläutert wurde, weshalb es sinnvoll ist, die Volatilität eines Assets selbst als einen stochastischen Prozess zu modellieren, stellt sich jedoch die Frage, auf welche Art und Weise dies geschehen soll. Dazu wäre es hilfreich, wenn die Volatilität  $\sigma_t$  des Assets tatsächlich als Zeitreihe vorläge. Hier besteht jedoch das Problem, dass die Volatilität eines Assets im Gegensatz zum Preis nicht direkt am Markt beobachtet werden kann, weshalb Methoden zur Schätzung dieser Zeitreihe verwendet werden müssen. Ein erster Ansatz, um die Volatilität eines Assets zu einem gegebenen Zeitpunkt zu bestimmen, besteht über die sogenannte implizite Volatilität. Hierbei wird in der Black-Scholes-Formel für den Preis einer Option die Volatilität so bestimmt, dass der theoretische Preis mit dem tatsächlich beobachteten Marktpreis einer Option auf das entsprechende Asset übereinstimmt. Zur Vollständigkeit sei hier noch einmal die Black-Scholes-Formel für den Preis eines europäischen Calls angegeben:

$$\begin{aligned} C(S, T) &= S\Phi(d_1) - e^{-rT}K\Phi(d_2) \\ d_1 &= \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \\ d_2 &= d_1 - \sigma\sqrt{T}. \end{aligned}$$

Das Auflösen der Black-Scholes-Formel nach der Volatilität  $\sigma$  ist möglich, da alle anderen Parameter, das heißt die Restlaufzeit  $T$ , der aktuelle Marktpreis  $S$  des Underlyings, der aktuelle Marktpreis der Option  $C(S, T)$ , der Basispreis  $K$  und der risikolose Zinssatz  $r$  am Markt beobachtbar sind. Die hieraus resultierende Volatilität wird als implizite Volatilität bezeichnet und kann beispielsweise mit Hilfe des Newton-Verfahrens numerisch berechnet werden. Hierbei tritt jedoch das Problem auf, dass diese für Optionen mit unterschiedlicher Laufzeit und unterschiedlichem Ausübungspreis verschieden ist. Typischerweise steigt sie umso mehr an, je stärker der Marktpreis vom Ausübungspreis entfernt ist. Wenn die implizite Volatilität für eine gegebene Laufzeit gegenüber dem Ausübungspreis grafisch dargestellt wird, so ergibt sich eine Kurve, die aufgrund ihres Aussehens auch als Volatility Smile bezeichnet wird. Wenn die Annahme der konstanten Volatilität innerhalb des Black-Scholes-Modells tatsächlich zuträfe, so müsste sich für unterschiedliche Optionen stets dieselbe implizite Volatilität ergeben. Insofern stützt das Vorhandensein des Volatility Smile erneut das Vorgehen, die Volatilität nicht mehr als eine konstante Größe anzunehmen. Wird die implizite Volatilität für verschiedene Ausübungspreise und Restlaufzeiten betrachtet, ergibt sich die sogenannte Volatilitätsoberfläche. Für eine beispielhafte Darstellung der Volatilitätsoberfläche, in der sich auch der Volatility Smile für verschiedene Laufzeiten erkennen lässt, sei auf [3, Figure 1.1] verwiesen. In dieser Abbildung wird die implizite Volatilität gegenüber der Log-Moneyness und der Restlaufzeit der Option dargestellt. Unter Moneyness wird hierbei der Quotient aus dem aktuellen Preis des Assets und dem Ausübungspreis der Option verstanden. Die Volatilitätsoberfläche ist von großer Bedeutung für die Kalibrierung von stochastischen Volatilitätsmodellen und wird daher auch im späteren Verlauf dieser Arbeit erneut von Relevanz sein.

### 3.3 Realisierte Volatilität

Da die implizite Volatilität aus den genannten Gründen nicht dazu geeignet ist, die Volatilität eines Assets zu bestimmen, wird im Folgenden ein alternatives Konzept zur Schätzung der Volatilität vorgestellt. Hierbei handelt es sich um die realisierte Volatilität. Für eine ausführliche Untersuchung der realisierten Volatilität, welche sich unter anderem auch mit dem Vorhandensein von Störungen des Preisprozesses auseinandersetzt, sei auf [13] verwiesen. In diesem Abschnitt wird vereinfachend angenommen, dass der Preisprozess eines Assets die SDE

$$dS_t = \mu_t S_t dt + \sigma_t S_t dB_t \quad (2)$$

erfüllt, wobei  $(B_t)$  eine Brownsche Bewegung ist und  $\mu_t, \sigma_t$  progressiv messbare, stochastische Prozesse seien mit  $\int_0^t (|\mu_s| + \sigma_s^2) ds < \infty$  fast sicher für alle  $t \geq 0$ . Damit ist der Preisprozess  $(S_t)$  insbesondere ein Ito-Prozess. Im Folgenden wird der Begriff der integrierten Varianz definiert, welche als die durchschnittliche Varianz von  $(S_t)$  auf einem gegebenem Intervall interpretiert werden kann.

**Definition 3.1** (Integrierte Varianz). *Sei  $(S_t)$  ein stochastischer Prozess, welcher die SDE (2) erfüllt. Dann ist die integrierte Varianz von  $S$  auf dem Intervall  $[t, t+\Delta]$  für  $t \geq 0$  und  $\Delta > 0$  definiert als*

$$IV_{t,\Delta} = \frac{1}{\Delta} \int_t^{t+\Delta} \sigma_s^2 ds.$$

In praktischen Anwendungen ist häufig die integrierte Varianz für einen Tag von Interesse, weshalb  $\Delta$  dann als die Zeitspanne eines Tages gewählt wird. In diesem Fall entspricht die integrierte Varianz der durchschnittlichen Varianz von  $S$  innerhalb eines Tages. Auch wenn einige sehr liquide Assets mit einer sehr hohen Frequenz gehandelt werden, so bildet sich der Marktpreis  $(S_t)$  doch stets nur zu endlich vielen Zeitpunkten. Die integrierte Varianz lässt sich somit ebenfalls nicht direkt am Markt beobachten. Es ist daher notwendig die integrierte Varianz auf Basis der beobachtbaren, endlichen Zahl an Realisierungen von  $(S_t)$  zu schätzen. Seien dazu nun für gegebene  $t \geq 0$  und  $\Delta > 0$  die Zeitpunkte zu denen der Preis  $(S_t)$  beobachtbar ist gegeben durch  $t = \tau_1 < \tau_2 < \dots < \tau_n = t + \Delta$ . Dann ist die realisierte Varianz wie folgt definiert.

**Definition 3.2** (Realisierte Varianz). *Sei  $\Lambda = (\tau_k)_{k=0,\dots,n} \in \mathcal{P}([t, t+\Delta])$  eine Partition von  $[t, t+\Delta]$  für  $t \geq 0$ ,  $\Delta > 0$  und  $n \in \mathbb{N}$ . Dann ist für einen Preisprozess  $(S_t)$  die realisierte Varianz definiert als*

$$RV_{t,\Delta}^\Lambda = \frac{1}{\Delta} \sum_{k=1}^n (\log(S_{\tau_k}) - \log(S_{\tau_{k-1}}))^2.$$

Die realisierte Varianz stellt eine Schätzung für die tatsächliche integrierte Varianz auf Basis von endlich vielen Realisierungen des Preisprozesses dar. Es soll nun gezeigt werden, dass die realisierte Varianz mit einer zunehmenden Anzahl an verfügbaren Realisierungen tatsächlich gegen die integrierte Varianz konvergiert. Dazu werden noch einmal der Begriff der quadratischen Variation wiederholt und die quadratische Variation eines Ito-Prozesses bestimmt.

**Definition 3.3** (Quadratische Variation). *Sei  $(X_t)$  ein reellwertiger stochastischer*



Prozess und  $\Lambda = (\tau_k)_{k=0,\dots,n} \in \mathcal{P}[0, t]$  eine Partition von  $[0, t]$  für ein  $t \geq 0$ . Dann ist die endliche quadratische Variation von  $X$  über  $\Lambda$  definiert als

$$Q_t^\Lambda(X) := \sum_{k=1}^n (X_{\tau_k} - X_{\tau_{k-1}})^2.$$

Die Maschengröße von  $\Lambda$  ist wie folgt definiert:

$$|\Lambda| := \max_{k=1,\dots,n} \{\tau_k - \tau_{k-1}\}.$$

Falls ein reellwertiger stochastischer Prozess  $\langle X, X \rangle$  existiert, so dass für jedes  $t > 0$  und jede Folge von Partitionen  $\Lambda_n \in \mathcal{P}([0, t])$  mit  $\lim_{n \rightarrow \infty} |\Lambda_n| = 0$  die Folge  $Q_t^{\Lambda_n}(X)$  in Wahrscheinlichkeit gegen  $\langle X, X \rangle_t$  konvergiert für  $n \rightarrow \infty$ , so heißt  $\langle X, X \rangle$  die quadratische Variation von  $X$ .

Wie sich die quadratische Variation eines Ito-Prozesses bestimmen lässt, ist Inhalt des folgenden Lemmas.

**Lemma 3.1.** Sei  $(X_t)$  ein Ito-Prozess mit

$$X_t = \int_0^t a_s ds + \int_0^t b_s dB_s,$$

dann gilt

$$\langle X, X \rangle_t = \int_0^t b_s^2 ds.$$

*Beweis.* Siehe [14, Theorem 3.26]. □

Aus dem vorangegangenen Lemma lässt sich schlussfolgern, dass die integrierte Varianz des Preisprozesses gerade der quadratischen Variation des Log-Preisprozesses entspricht.

**Theorem 3.1.** Sei ein Preisprozess  $(S_t)$  gegeben, welcher die SDE (2) erfüllt. Dann gilt

$$\Delta \cdot IV_{t,\Delta} = \langle \log(S), \log(S) \rangle_{t+\Delta} - \langle \log(S), \log(S) \rangle_t$$

für alle  $t, \Delta \geq 0$ . Insbesondere gilt  $RV_{t,\Delta}^{\Lambda_n} \rightarrow IV_{t,\Delta}$  für  $n \rightarrow \infty$  in Wahrscheinlichkeit für jede Folge von Partitionen  $\Lambda_n$  mit  $\lim_{n \rightarrow \infty} |\Lambda_n| = 0$ .

*Beweis.* Mit der Ito-Formel für Ito-Prozesse folgt für alle  $t \geq 0$ :

$$\begin{aligned} d\log(S_t) &= \frac{1}{S_t} dS_t - \frac{1}{2S_t^2} d\langle S, S \rangle_t \\ &= \mu_t dt + \sigma_t dB_t - \frac{1}{2} \sigma_t^2 dt \\ &= \left( \mu_t - \frac{1}{2} \sigma_t^2 \right) dt + \sigma_t dB_t. \end{aligned}$$

Somit ist  $\log(S_t)$  ebenfalls ein Ito-Prozess und es gilt mit Lemma 3.1

$$\langle \log(S), \log(S) \rangle_t = \int_0^t \sigma_s^2 ds$$

und damit insbesondere

$$\langle \log(S), \log(S) \rangle_{t+\Delta} - \langle \log(S), \log(S) \rangle_t = \int_t^{t+\Delta} \sigma_s^2 ds = \Delta \cdot IV_{t,\Delta}.$$

Dass  $RV_{t,\Delta}^{\Lambda_n}$  in Wahrscheinlichkeit gegen  $IV_{t,\Delta}$  konvergiert folgt jetzt aus Definition 3.3, da  $\Delta \cdot RV_{t,\Delta}^{\Lambda} = Q_{t+\Delta}^{\Lambda}(\log(S)) - Q_t^{\Lambda}(\log(S))$ .  $\square$

Theorem 3.1 besagt nun also, dass die realisierte Varianz in Wahrscheinlichkeit gegen die integrierte Varianz konvergiert, wenn die Maschengröße gegen 0 konvergiert. Somit lässt sich die durchschnittliche tägliche Varianz eines Assets schätzen, indem die Summe der quadrierten Log-Renditen innerhalb eines Tages gebildet wird. Dies liefert nun eine Möglichkeit, um den Volatilitätsprozess eines Assets zu untersuchen, indem die Zeitreihe der täglichen realisierten Volatilität betrachtet wird. Dieses Verfahren macht damit den Volatilitätsprozess eines Assets beobachtbar. Unter der realisierten Volatilität wird im Folgenden die radizierte realisierte Varianz verstanden.

Es sei angemerkt, dass die realisierte Varianz mit zunehmender Frequenz der Preisdaten gegen die integrierte Varianz konvergiert, jedoch nicht klar ist, mit welcher Geschwindigkeit dies geschieht. So zeigten beispielsweise Barndorff-Nielsen und Shephard in [15], dass insbesondere in Zeiten von hoher Volatilität die tägliche realisierte Varianz ein großes Konfidenzintervall besitzt und damit im Zweifel deutlich von der tatsächlichen integrierten Varianz abweichen kann. Ein weiteres Problem ist sogenanntes Market Microstructure Noise, welches dafür sorgen kann, dass die realisierte Varianz mit zunehmender Anzahl an Realisierungen nicht gegen die integrierte Varianz konvergiert. So zeigten beispielsweise Bandi und Russell in [16]

dass die realisierte Varianz bei Vorhandensein von Market Microstructure Noise unter bestimmten Bedingungen mit zunehmender Anzahl an Realisierungen divergiert.

Nachdem das Konzept der realisierten Volatilität als Schätzung der tatsächlichen, instantanten Volatilität eingeführt wurde, werden im kommenden Abschnitt die realisierten Volatilitätsprozesse von zwei verschiedenen Datensätzen mit historischen Marktdaten untersucht werden. Es wird sich dabei herausstellen, dass der realisierte Volatilitätsprozess sehr gut mit Hilfe einer fraktalen Brownschen Bewegung modelliert werden kann.

## 4 Volatilität ist rau

In ihrem Artikel „Volatility is rough“ [3] zeigten Gatheral, Jaisson und Rosenbaum, dass sich der Volatilitätsprozess von vier verschiedenen Assets (DAX-Futures, BUND-Futures, S&P 500 und NASDAQ Index) mit Hilfe einer fraktalen Brownsche Bewegung modellieren lässt. Ziel dieses Kapitels ist es, diese Ergebnisse anhand von zwei Datensätzen zu verifizieren. Dazu wird die realisierte Volatilität des S&P 500 Index, welcher die nach Marktkapitalisierung 500 größten US-amerikanischen Unternehmen enthält, und der Microsoft-Aktie über einen mehrjährigen Zeitraum untersucht. Während in [3] lediglich Aktienindizes untersucht wurden, wird in diesem Abschnitt auch die Volatilität einer Einzelaktie betrachtet. Die fraktale Brownsche Bewegung ist dadurch gekennzeichnet, dass ihre Inkremente eine spezielle Skalierungseigenschaft aufweisen, die vom Hurst-Parameter  $H$  abhängt. Im folgenden Abschnitt wird gezeigt, dass die logarithmierte realisierte Volatilität genau dieselbe Skalierungseigenschaft wie eine fraktale Brownsche Bewegung aufweist. Im Rahmen dieser Untersuchung wird dann auch der Hurst-Parameter der logarithmierten realisierten Volatilität geschätzt. Auch wenn diese Ergebnisse darauf schließen lassen, dass die Log-Volatilität durch eine fraktale Brownsche Bewegung beschrieben werden kann, zeigt eine genauere Untersuchung der Inkremente, dass große Wertänderungen, im Vergleich zur Normalverteilung, überdurchschnittlich häufig auftreten.

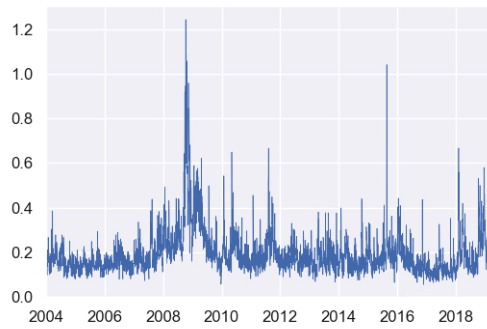
### 4.1 Vorstellung der Datensätze

In der vorliegenden empirischen Untersuchung werden zwei Datensätze betrachtet. Der erste Datensatz enthält die realisierte Varianz der Microsoft-Aktie aus dem Zeitraum vom 01.02.2004 bis zum 15.04.2019 was einer Gesamtanzahl von 3487 Handelstagen entspricht. Die realisierte Varianz wurde auf Basis der nationalen, niedrigsten Ask- und höchsten Bid-Preise bestimmt, welche von Tick Data bereitgestellt werden.<sup>1</sup> Der zweite verwendete Datensatz enthält die realisierte Varianz des S&P 500 Index im Zeitraum vom 03.01.2000 bis zum 01.06.2020 was einer Gesamtanzahl von 5120 Handelstagen entspricht und entstammt der frei zugänglichen Oxford-Man Realized Library.<sup>2</sup> In beiden Datensätzen wurde die realisierte Varianz auf Basis von 5-minütigen Renditen berechnet.

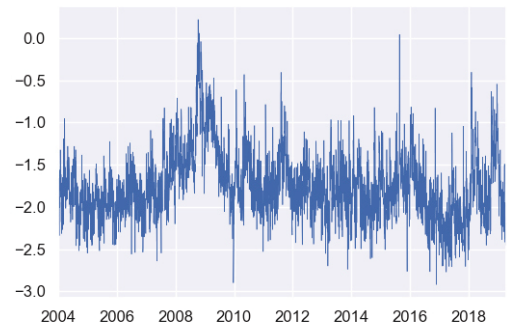
---

<sup>1</sup><https://www.tickdata.com/product/nbbo>

<sup>2</sup><https://realized.oxford-man.ox.ac.uk/data>

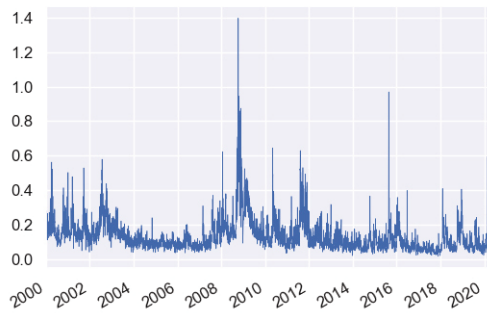


(a) realisierte Volatilität



(b) logarithmierte realisierte Volatilität

Abbildung 4: (logarithmierte) realisierte Volatilität der Microsoft-Aktie

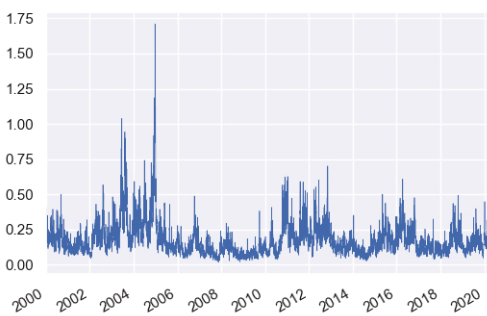


(a) realisierte Volatilität



(b) logarithmierte realisierte Volatilität

Abbildung 5: (logarithmierte) realisierte Volatilität des S&P 500



(a) realisierte Volatilität



(b) logarithmierte realisierte Volatilität

Abbildung 6: simulierte (logarithmierte) realisierte Volatilität

In den Abbildungen (4), (5) und (6) sind die (logarithmierten) Volatilitätsprozesse des S&P 500 und der Microsoft-Aktie einem simulierten Volatilitätsprozess gegenübergestellt. Der simulierte Log-Volatilitätsprozess ist durch die Simulation einer

skalierten fraktalen Brownschen Bewegung entstanden, wobei sich die Parameter eben jener aus den Parameter-Schätzungen für den S&P 500 ergeben. Wie genau diese Schätzungen zustande kommen, wird in den folgenden Abschnitten erläutert. Es ist jedoch bereits zu erkennen, dass die logarithmierten Volatilitätsprozesse beider Assets sehr stark einer fraktalen Brownschen Bewegung ähneln. Der simulierte (Log-)Volatilitätsprozess lässt sich mit bloßem Auge nicht von den tatsächlichen Volatilitätsprozessen unterscheiden. Die Pfade der logarithmierten Volatilitätsprozesse weisen insbesondere ein antipersistentes Verhalten auf, das heißt Phasen von steigender Volatilität sind häufig gefolgt von Phasen niedriger Volatilität und umgekehrt. Dies ist auch ein bekanntes Verhalten einer fraktalen Brownschen Bewegung mit Hurst-Parameter  $H < 0.5$ . Wie im folgenden Abschnitt gezeigt werden wird, lassen sich beide Log-Volatilitätsprozesse gut durch eine fraktale Brownsche Bewegung mit  $H \approx 0.1$  modellieren. Ein so niedriger Hurst-Parameter spiegelt die Tatsache wieder, dass die Pfade der Log-Volatilität ein stark antipersistentes Verhalten aufweisen und deshalb sehr rau erscheinen.

## 4.2 Die Skalierungseigenschaft der Log-Volatilität

Ziel dieses Unterabschnittes ist es, zu zeigen, dass die logarithmierte realisierte Volatilität sowohl des S&P 500 Index als auch der Microsoft-Aktie dieselbe Skalierungseigenschaft wie die eine fraktale Brownsche Bewegung aufweist. Die Ausführungen zur Bestimmung des Hurst-Parameters  $H$  folgen dabei [3]. Wie im Kapitel über die grundlegenden Eigenschaften der fraktalen Brownschen Bewegung gezeigt wurde, erfüllen die Inkremente dieses stochastischen Prozesses die folgende Eigenschaft:

$$\mathbb{E} [|B_t^H - B_s^H|^q] = K_q |t - s|^{qH}.$$

Hierbei ist  $K_q = \mathbb{E} [|B_1^H|^q]$  das  $q$ -te absolute Moment einer standardnormalverteilten Zufallsvariable. Es wird nun angenommen, dass sich der Log-Volatilitätsprozess  $\log(\sigma_t)$  wie eine skalierte fraktale Brownschen Bewegung verhält, das heißt  $\log(\sigma_t) - \log(\sigma_s) = \nu (B_t^H - B_s^H)$  für ein  $\nu > 0$ . In diesem Fall gelte

$$\mathbb{E} [|\log(\sigma_{t+\Delta}) - \log(\sigma_t)|^q] = \nu^q K_q \Delta^{qH}.$$

Durch Logarithmieren auf beiden Seiten ergibt sich damit

$$\log(\mathbb{E} [|\log(\sigma_{t+\Delta}) - \log(\sigma_t)|^q]) = \log(K_q) + q \log(\nu) + qH \log(\Delta). \quad (3)$$

Bei Gleichung (3) handelt es sich offensichtlich um eine lineare Funktion in  $\log(\Delta)$ . Der Anstieg dieser Funktion ist gegeben durch  $qH$ , was bedeutet, dass dieser mit wachsendem  $q$  proportional mit dem Faktor  $H$  ansteigt. Es gilt zu prüfen, ob die Zeitreihe der realisierten Volatilität genau diese Eigenschaft aufweist.

#### 4.2.1 Schätzung des Hurst-Parameters

Seien  $N$  Realisierungen  $\sigma_0, \sigma_\Delta, \dots, \sigma_{N\Delta}$  eines Volatilitätsprozesses gegeben. Nun definiere für  $q > 0$

$$m(q, \Delta) = \frac{1}{N} \sum_{k=1}^N |\log(\sigma_{k\Delta}) - \log(\sigma_{(k-1)\Delta})|^q,$$

was unter der Bedingung, dass die Inkremente stationär sind und das starke Gesetz der großen Zahlen gilt, eine Schätzung für

$$\mathbb{E} [|\log(\sigma_{t+\Delta}) - \log(\sigma_t)|^q] = \mathbb{E} [|\log(\sigma_\Delta) - \log(\sigma_0)|^q]$$

darstellt. Falls der logarithmierte realisierte Volatilitätsprozess sich tatsächlich in etwa wie eine fraktale Brownsche Bewegung verhalten sollte, wäre wegen Gleichung (3) zu erwarten, dass die Abbildung  $\log(\Delta) \rightarrow \log(m(q, \Delta))$  für verschiedene  $q$  jeweils linear verläuft.

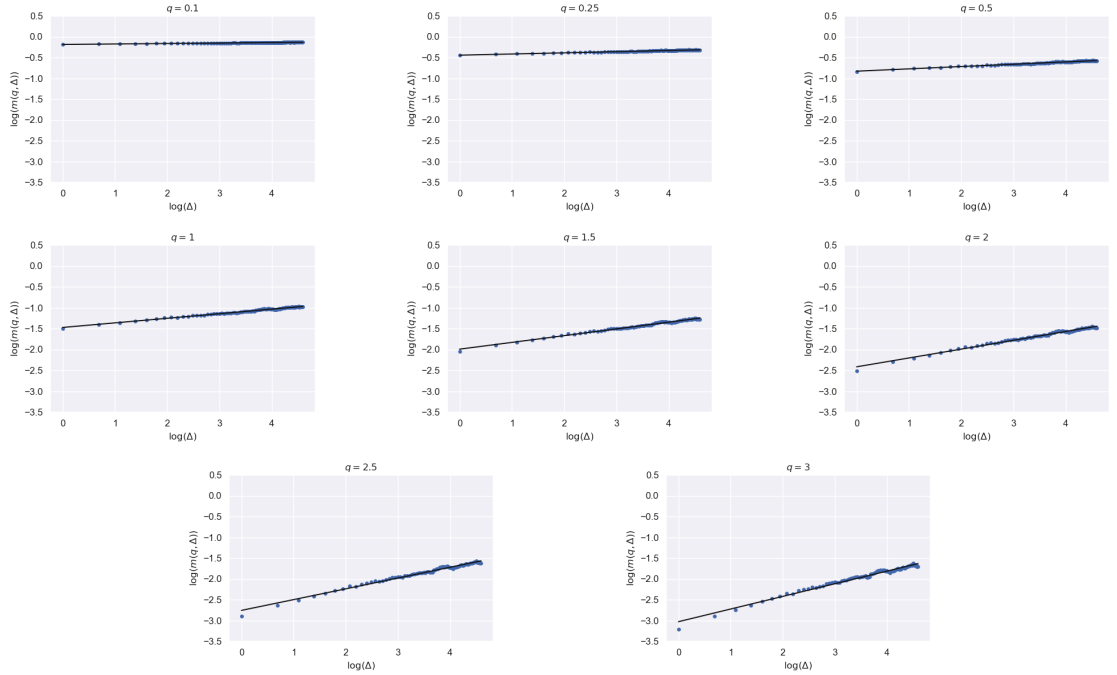


Abbildung 7: Plot von  $\log(\Delta) \rightarrow \log(m(q, \Delta))$  für verschiedene  $q$  (MSFT)

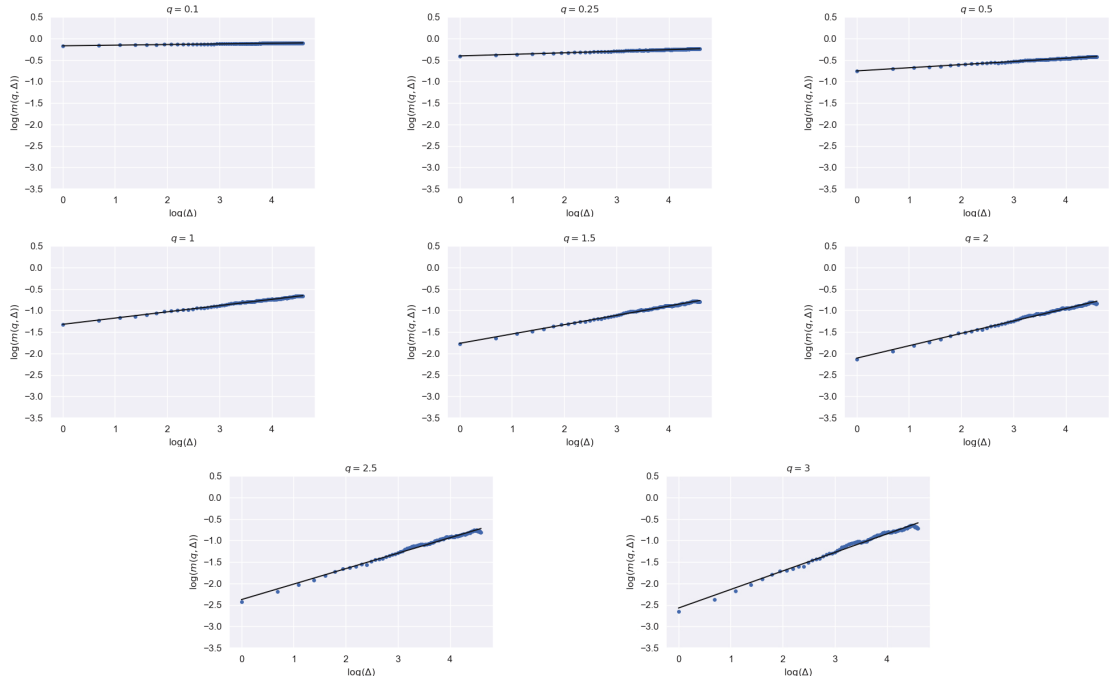


Abbildung 8: Plot von  $\log(\Delta) \rightarrow \log(m(q, \Delta))$  für verschiedene  $q$  (S&P 500)

In Abbildung 7 und Abbildung 8 ist zu sehen, dass für verschiedene  $q$  die Werte von  $\log(m(q, \Delta))$  sowohl für die Microsoft-Aktie als auch den S&P 500 Index tatsächlich in etwa auf einer Geraden liegen. Für  $\Delta$  wurden jeweils die Werte  $\Delta = 1, 2, \dots, 100$  verwendet. Um eine Gerade zu bestimmen, die in etwa durch die gegebenen Punkte verläuft, wurde für jedes  $q$  jeweils eine lineare Regression durchgeführt. Die Abweichungen von den Geraden sind für alle  $q$  in einem Bereich, der sich auch bei einer simulierten fraktalen Brownschen Bewegung beobachten lässt. Der Anstieg der jeweiligen Geraden für jedes  $q$  wird im Folgenden mit  $s_q$  und das Absolutglied mit  $a_q$  bezeichnet. Es ist zu erkennen, dass mit steigendem  $q$  der Anstieg  $s_q$  zu- und das Absolutglied  $a_q$  abnimmt. Für eine fraktale Brownsche wäre der Anstieg von  $s_q$  wegen Gleichung (3) proportional zu  $q$ . In Abbildung 9 ist zu sehen, dass auch für die beiden Datensätze der Anstieg  $s_q$  proportional zu  $q$  verläuft, was daran zu erkennen ist, dass die Datenpunkte ziemlich genau auf einer Geraden liegen, die durch den Ursprung verläuft. Die Geraden durch die jeweiligen Datenpunkte wurden erneut mit Hilfe einer linearen Regression, jedoch ohne Absolutglied, bestimmt. Die logarithmierte realisierte Volatilität weist also für beide Datensätze in etwa die gleiche Skalierungseigenschaft auf, wie eine fraktale Brownsche Bewegung. Der Anstieg der Geraden aus Abbildung 9 liefert dann eine Schätzung den Hurst-Parameter  $H$ . In Abbildung 9 ist bereits zu erkennen, dass der Hurst-Parameter für die Microsoft-Aktie geringer ist als für den S&P 500 Index, da die entsprechende Gerade einen geringeren Anstieg aufweist. Genauer sind die Schätzungen der Hurst-Parameter für



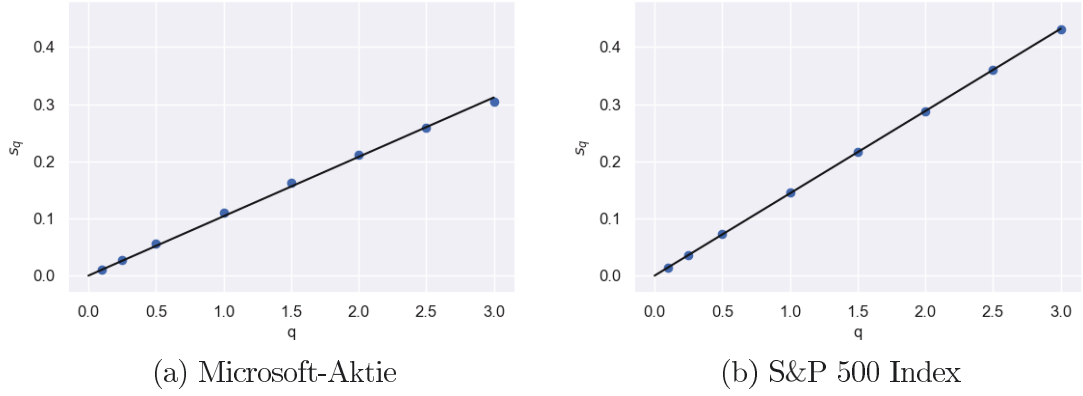


Abbildung 9: Plot von  $q \rightarrow s_q$

beide Datensätze gegeben durch

$$H_{MSFT}^{est} \approx 0.104$$

$$H_{S\&P}^{est} \approx 0.144.$$

Der hier auf Basis einer längeren Zeitreihe geschätzte Hurst-Parameter für die Log-Volatilität des S&P 500 entspricht ziemlich genau jenem, den Gatheral et. al in [3] beobachten konnten, auch wenn ein längerer Beobachtungszeitraum vorliegt.

#### 4.2.2 Schätzung des Skalierungsfaktors

Es stellt sich noch die Frage, wie für beide Datensätze die Volatilität der Volatilität, also der Skalierungsfaktor  $\nu$ , bestimmt werden kann. Aus Gleichung (3) lässt sich ablesen, dass im Falle einer fraktalen Brownschen Bewegung die Absolutglieder  $a_q$  in etwa dem Ausdruck  $\log(K_q) + q \log(\nu)$  entsprechen, das heißt  $a_q - \log(K_q)$  sollte proportional zu  $q$  sein. In Abbildung 10 ist zu erkennen, dass auch dies für beide Datensätze ziemlich exakt erfüllt ist. Alle Datenpunkte liegen in etwa auf einer Geraden, welche durch den Ursprung verläuft. Der Anstieg dieser Geraden entspricht dann dem Proportionalitätsfaktor  $\log(\nu)$  und liefert damit eine Schätzung für den Skalierungsfaktor  $\nu$ , welche auch als Volatilität der Volatilität bezeichnet werden kann. Für die beiden Datensätze gelten die folgenden Schätzungen:

$$\nu_{MSFT}^{est} \approx 0.304$$

$$\nu_{S\&P}^{est} \approx 0.354.$$

Wie im vorangegangenen Abschnitt dargestellt, wurde der Volatilitätsprozess mit Hilfe der realisierten Volatilität geschätzt, das heißt statt der tatsächlichen instan-

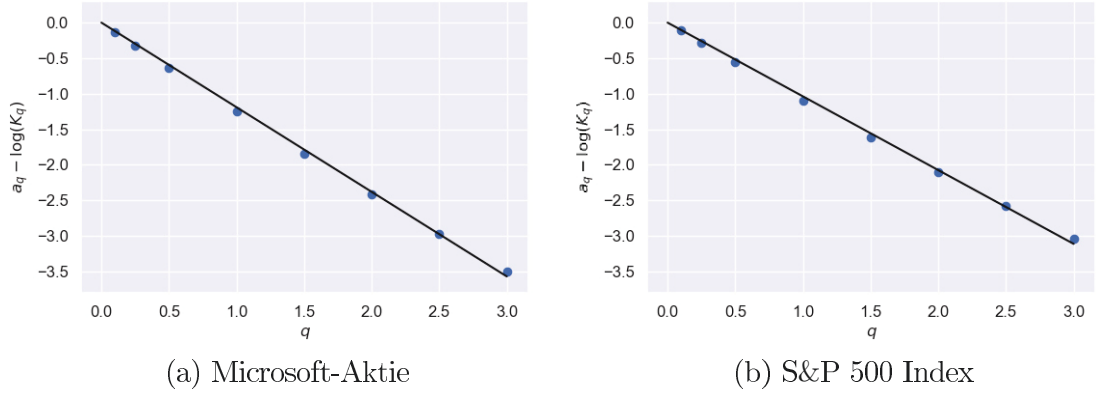


Abbildung 10: Plot von  $q \rightarrow a_q - \log(K_q)$

tanen Volatilität wird die durchschnittliche Volatilität innerhalb eines Tages betrachtet. Diese Durchschnittsbildung führt jedoch zu einer gewissen Glättung des tatsächlichen Volatilitätsprozesses. Wie in [3, Abschnitt 3.4] mit Hilfe von Simulationen gezeigt wurde, führt diese Glättung dazu, dass aufgrund der Durchschnittsbildung der tatsächliche Hurst-Parameter  $H$  über- und der Skalierungsparameter  $\nu$  unterschätzt werden. Es ist also zu erwarten, dass der tatsächliche Volatilitätsprozess noch rauere Pfade als die realisierte Volatilität aufweist.

### 4.3 Untersuchung der Inkremente

Nachdem gezeigt wurde, dass die logarithmierte realisierte Volatilität in etwa die gleiche Skalierungseigenschaft aufweist, wie eine fraktale Brownsche Bewegung, sollen nun die Inkremente der Log-Volatilität genauer untersucht werden. Für eine fraktale Brownsche Bewegung  $B_t^H$  sind die Inkremente stets normalverteilt, genauer gilt

$$B_{t+\Delta}^H - B_t^H \sim \mathcal{N}(0, \Delta^{2H})$$

für alle  $t, \Delta > 0$ . Sollte die Log-Volatilität einer fraktalen Brownschen Bewegung folgen, ist zu erwarten, dass die Inkremente für verschiedene  $\Delta$  stets normalverteilt sind, also

$$\log(\sigma_{t+\Delta}) - \log(\sigma_t) \sim \mathcal{N}(0, \nu^2 \Delta^{2H}). \quad (4)$$

In den Abbildungen 11 und 12 sind die Histogramme der Inkremente der logarithmierten realisierten Volatilität für verschiedene  $\Delta$  dargestellt. Weiterhin ist die Dichte der Normalverteilung zu sehen, die sich jeweils durch eine Schätzung von Er-

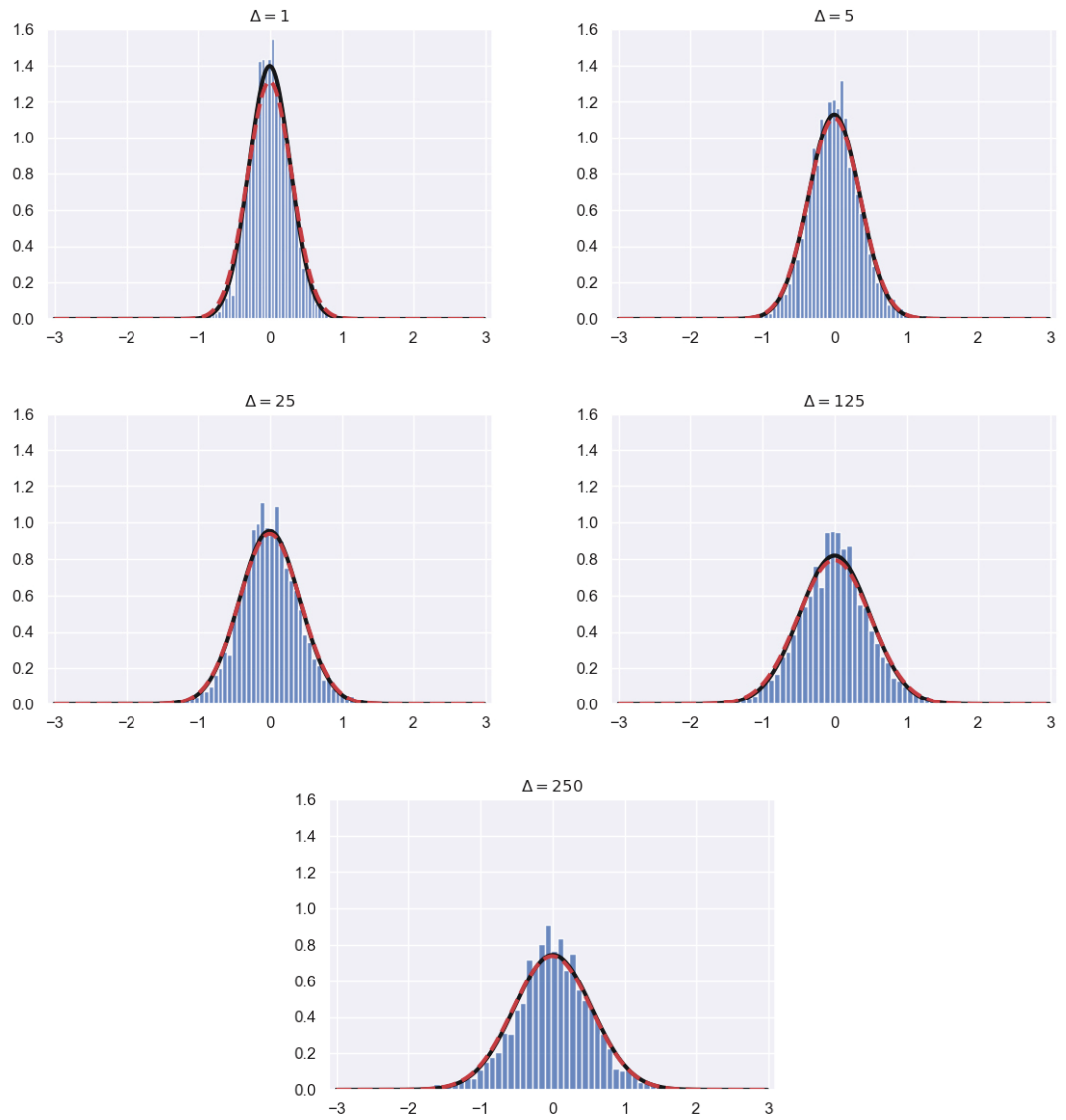


Abbildung 11: Histogramme der Inkremente der logarithmierten realisierten Volatilität (MSFT),  $\Delta$  in Tagen

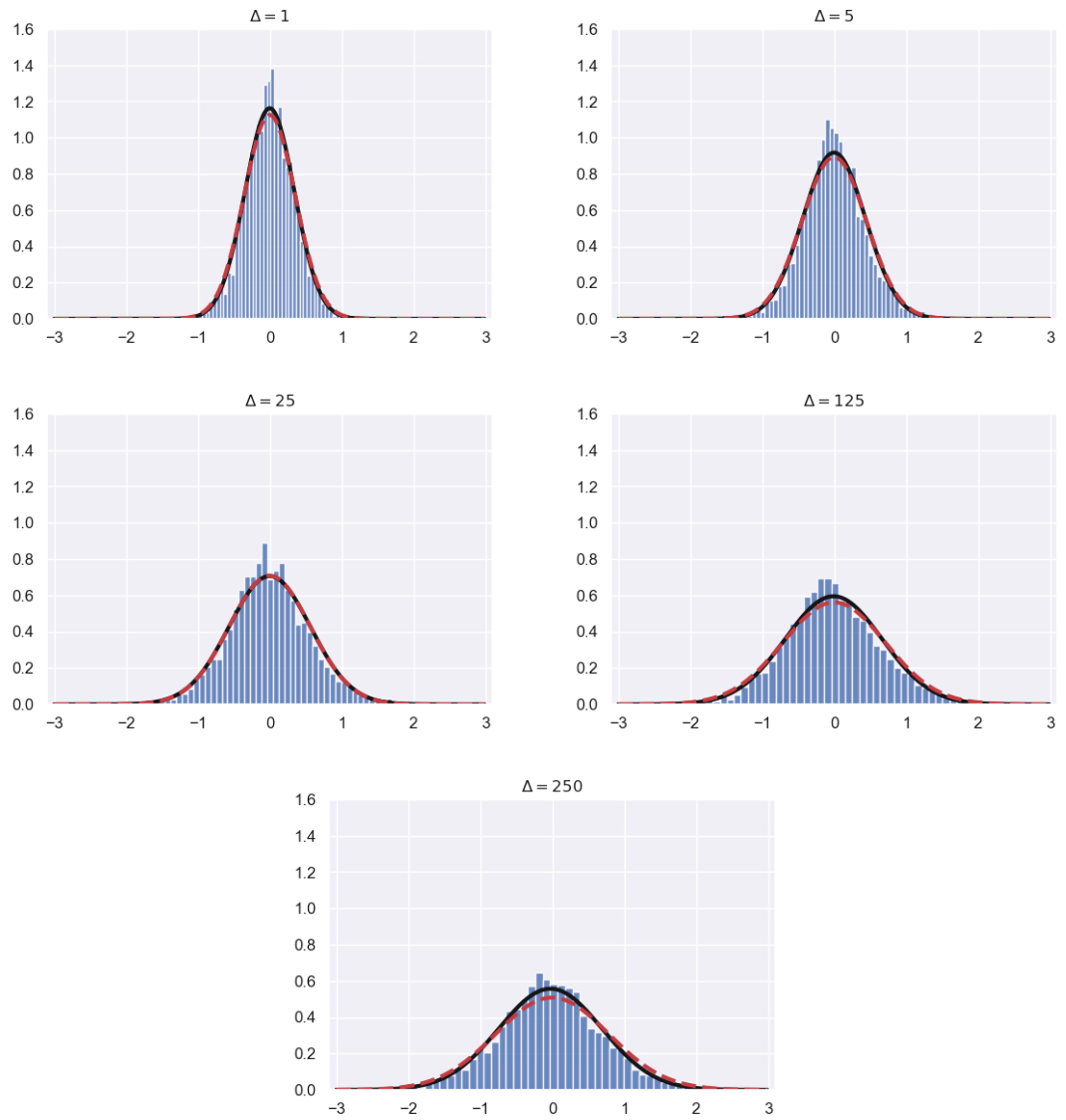


Abbildung 12: Histogramme der Inkremente der logarithmierten realisierten Volatilität (S&P 500),  $\Delta$  in Tagen

wartungswert und Standardabweichung ergibt (schwarze Linie). Die rote, gestrichelte Linie ist jeweils die Dichte der Normalverteilung, die sich aus Gleichung (4) ergibt, wobei für  $H$  und  $\nu$  die Schätzungen aus dem vorangegangenen Unterabschnitt verwendet wurden. Es ist zu sehen, dass die rote und schwarze Linie immer nur marginal voneinander abweichen, was erneut die These stützt, dass die Log-Volatilität dieselbe Skalierungseigenschaft wie eine fraktale Brownsche Bewegung aufweist. Auch ist die Abweichung der Histogramme von der Normalverteilungsdichte nicht allzu groß. Jedoch lässt sich für alle  $\Delta$  erkennen, dass die Inkremente tendenziell eher spitzer verteilt sind als es eine Normalverteilung implizieren würde. Auch an den Rändern besitzen die Verteilungen der Inkremente, im Vergleich zu einer Normalverteilung, mehr Gewicht. Laut [3] bewegten sich diese Abweichungen jedoch im Rahmen dessen, was sich auch bei einer simulierten fraktalen Brownschen Bewegung derselben Größe beobachten ließe. Dies kann auf Basis der vorliegenden Daten jedoch nicht bestätigt werden, wie die folgenden Ausführungen deutlich machen.

Eine Möglichkeit die Verteilung der Inkremente im Vergleich zu einer Normalverteilung genauer zu untersuchen ist ein sogenannter q-q-Plot. Hierbei werden die empirischen Quantile des Datensatzes mit den theoretischen Quantilen einer Normalverteilung verglichen. Erwartungswert und Standardabweichung der Normalverteilung ergeben sich durch eine Maximum-Likelihood Schätzung auf den gegebenen Daten. In den Abbildungen 13 und 14 sind die q-q-Plots der Inkremente der logarithmierten realisierten Volatilität für die Microsoft-Aktie und den S&P 500 Index zu sehen. Es ist zu erkennen, dass für beide Datensätze und alle Inkrementgrößen  $\Delta$  die Quantile deutlich von der Normalverteilung abweichen. Dies wird dadurch deutlich, dass die Datenpunkte von der Diagonalen abweichen. Wären die Daten perfekt normalverteilt, so lägen alle Punkte auf dieser Geraden. Insbesondere extrem große Werte treten in beiden Datensätzen deutlich häufiger auf, als es bei einer Normalverteilung zu erwarten wäre. Dies ist daran zu erkennen, dass die sehr hohen Quantile des Datensatzes deutlich größer sind, als die theoretischen Quantile der Normalverteilung. Auch für extrem niedrige Werte lässt sich dies beobachten jedoch in einem geringeren Maße als für die hohen Werte. Hier ist die Abweichung von der Diagonalen deutlich geringer, tritt auch erst für größere Quantile auf und liegt noch im Rahmen dessen, was sich auch bei einer simulierten fraktalen Brownschen Bewegung beobachten ließe. Dass für sehr große Werte das beobachtete Quantil jedoch systematisch über dem theoretischen Quantil liegt spricht dafür, dass es sich hier nicht um eine Abweichung handelt, die lediglich auf die zufällige Natur der Daten zurückzuführen ist. Wären die Daten tatsächlich normalverteilt, so wäre es genau so

wahrscheinlich, dass das beobachtete Quantil unter dem theoretischen Quantil liegt. Jedoch ist für alle  $\Delta$  und beide Datensätze zu beobachten, dass für extrem große Werte stets das beobachtete über dem theoretischen Quantil liegt.

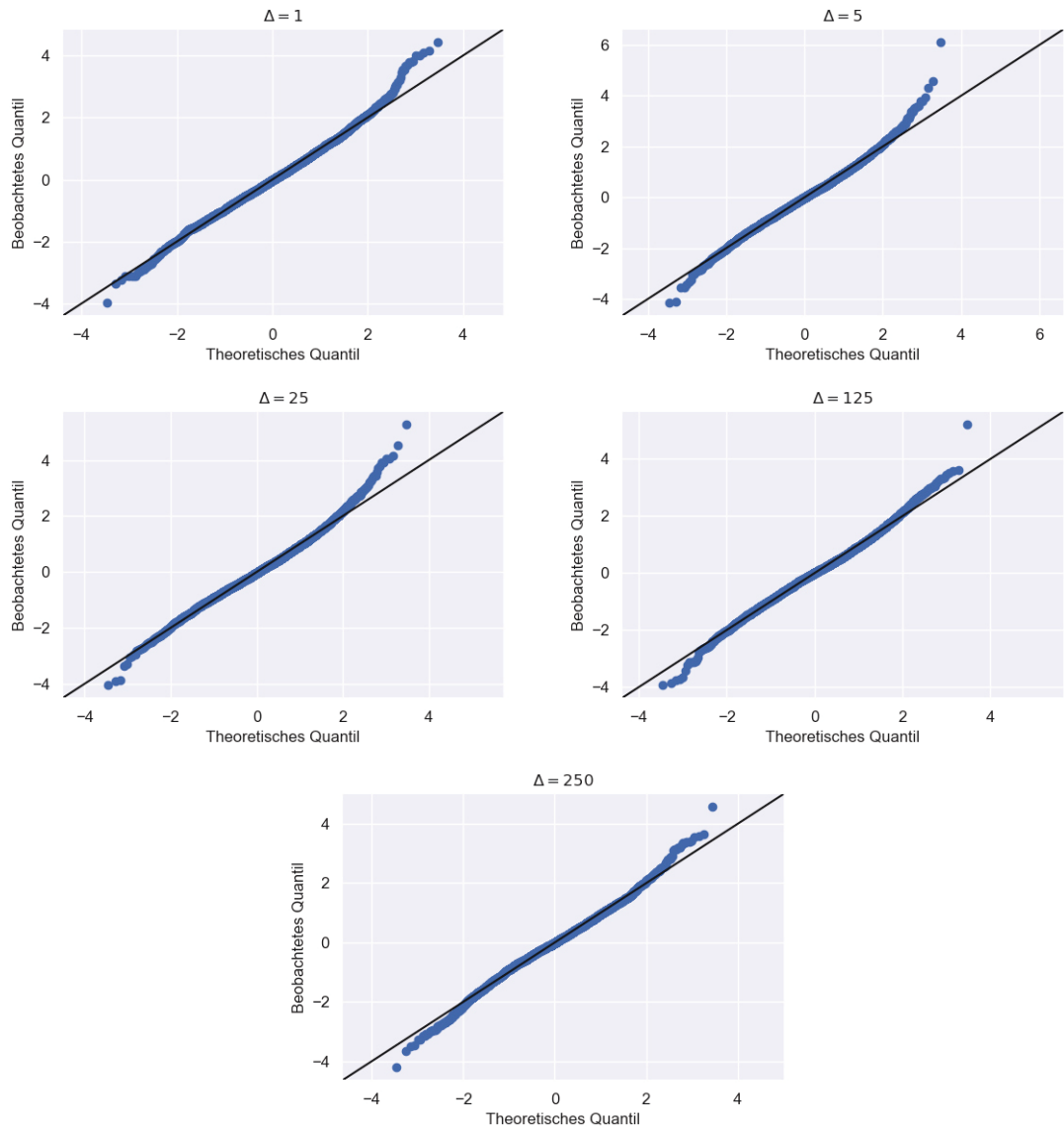


Abbildung 13: q-q-Plot der Inkremente der logarithmierten realisierten Volatilität (MSFT),  $\Delta$  in Tagen

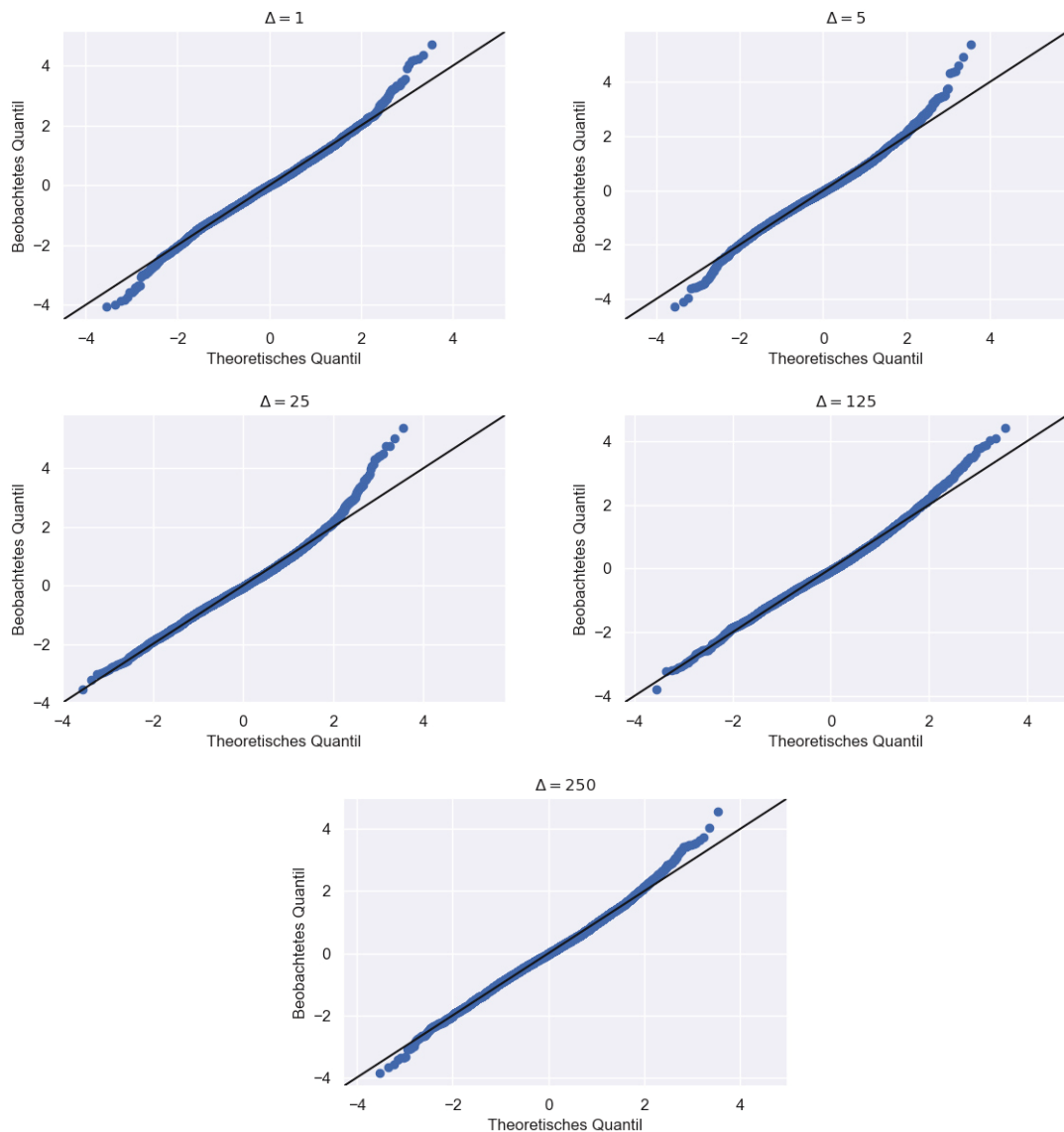


Abbildung 14: q-q-Plot der Inkremente der logarithmierten realisierten Volatilität (S&P 500),  $\Delta$  in Tagen

Einen weiteren Hinweis darauf, wie die Daten verteilt sind, liefern die Schiefe und die Kurtosis. Bei der Schiefe (resp. Kurtosis) handelt es sich um das zentrierte, standardisierte Moment dritter (resp. vierter) Ordnung. Die Schiefe der Daten liefert ein Maß dafür, wie asymmetrisch die Daten verteilt sind. Für eine Standardnormalverteilung ist die Schiefe gleich Null. Eine negative Schiefe bedeutet, dass die Dichtefunktion links vom Erwartungswert flacher abfällt als rechts und extrem niedrige Werte häufiger auftreten. Für eine positive Schiefe hingegen fällt die Dichte nach rechts flacher ab und extrem positive Werte treten gehäuft auf. Die Kurtosis ist ein Maß für dafür, wie spitz die Dichtefunktion einer Zufallsvariablen ist. Für eine Normalverteilung hat die Kurtosis den Wert 3. Liegt die Kurtosis über (resp.

unter) drei, so ist die Dichte spitzer (resp. flacher) als bei einer Normalverteilung und extreme Werte treten im Vergleich häufiger (resp. seltener) auf (vgl. [17, Kapitel 19.2.2]).

$\Delta$	Kurtosis	Skewness
1	3.694	0.194
5	4.066	0.175
25	4.033	0.274
125	3.784	0.09
250	3.784	-0.017

(a) MSFT

$\Delta$	Kurtosis	Schiefe
1	3.826	0.084
5	4.221	0.196
25	4.238	0.422
125	3.627	0.308
250	3.566	0.143

(b) S&P 500

Abbildung 15: Kurtosis und Schiefe der Inkremente,  $\Delta$  in Tagen

In den Tabellen in Abbildung 15 befinden sich die Schätzungen der Schiefe und Kurtosis für die Inkremente der logarithmierten realisierten Volatilität. Es ist deutlich zu sehen, dass die Inkremente für alle  $\Delta$  eine Kurtosis größer als drei und eine positive Schiefe aufweisen. Dies macht noch einmal deutlich, dass extreme Werte und insbesondere extrem große Werte im Vergleich zu einer Normalverteilung häufiger auftreten. Mit Hilfe der Kurtosis und Schiefe lässt sich ebenfalls ein statistischer Test auf Normalverteilung durchführen, der sogenannte Jarque-Bera-Test. Hierbei wird für eine Schiefe  $S$  und eine Kurtosis  $K$  bei einem Datensatz der Größe  $N$  die folgende Teststatistik verwendet:

$$JB = N \left( \frac{S^2}{6} + \frac{(K - 3)^2}{24} \right).$$

Jarque und Bera zeigten in [18], dass für den Fall, dass die zugrundeliegenden Daten einer Normalverteilung entstammen, die Teststatistik stets einer  $\chi^2_{(2)}$ -Verteilung folgt. Die Nullhypothese, dass die Daten normalverteilt sind, kann daher dann abgelehnt werden, wenn  $JB$  für ein vorgegebenes Signifikanzniveau  $\alpha$  den kritischen Wert der  $\chi^2_{(2)}$ -Verteilung übersteigt. Für ein Signifikanzniveau von  $\alpha = 0.05$  liegt dieser kritische Wert beispielsweise bei 5.99. In Abbildung 16 ist zu sehen, dass dieser kritische Wert für beide Datensätze deutlich überschritten ist. Die Normalverteilungsannahme muss somit, bei einem Signifikanzniveau von  $\alpha = 0.05$ , für beide Datensätze abgelehnt werden.

Zusammenfassend lässt sich also sagen, dass für beide Datensätze die logarithmierte realisierte Volatilität ziemlich genau die gleiche Skalierungseigenschaft wie eine fraktale Brownsche Bewegung aufweist. Für beide Datensätze liegt der geschätzte Hurst-Parameter in etwa bei 0.1, was deutlich unter 0.5 liegt. Daraus lässt sich schlie-



$\Delta$	$JB$
1	101
5	201
25	218
125	100
250	92

(a) MSFT

$\Delta$	$JB$
1	152
5	351
25	477
125	161
250	82

(b) S&P 500

Abbildung 16: Teststatistik  $JB$ ,  $\Delta$  in Tagen

ßen, dass die Log-Volatilität ein stark antipersistentes Verhalten aufweist. Auf eine Phase von steigender Volatilität folgt somit häufig eine Phase von fallender Volatilität. Diese Ergebnisse legen nahe, dass sich die Log-Volatilität sehr gut mit Hilfe einer skalierten fraktalen Brownschen Bewegung modellieren lässt. Unter dieser Annahme sollte zu beobachten sein, dass die Inkremente der Log-Volatilität normalverteilt sind. Ein genauerer Blick auf die Verteilung der Inkremente zeigt jedoch, dass im Vergleich zu einer Normalverteilung sehr große Werte deutlich häufiger auftreten, was nicht, wie in [3], mit der Zufälligkeit der Daten erklärt werden kann, da dies systematisch und unabhängig von der Größe der Inkremente zu beobachten ist. Auch der Jarque-Bera-Test bestätigt die These, dass die Inkremente nicht perfekt normalverteilt sind.

Nichtsdestotrotz wird im folgenden angenommen werden, dass die Log-Volatilität einer skalierten fraktalen Brownschen Bewegung entspricht. Im folgenden Abschnitt wird untersucht werden, wie sich diese Annahme auf die Bepreisung von Optionen auswirkt. Es wird gezeigt, wie sich die Log-Volatilität und der Preisprozess unter einem äquivalenten Martingalmaß verhalten und wie sich das daraus ergebende Modell simulieren und mit Hilfe von Marktdaten kalibrieren lässt. Dabei wird auch ein Ansatz zur Modellkalibrierung mit Hilfe künstlicher neuronaler Netze vorgestellt, welcher die Modellkalibrierung drastisch beschleunigen kann.

## 5 Risikoneutrale Bewertung unter rauer Volatilität

Im vorangegangenen Abschnitt wurde anhand von zwei Datensätzen gezeigt, dass sich die logarithmierte realisierte Volatilität, eine Schätzung für den Volatilitätsprozess  $\sigma_t$ , in etwa wie eine skalierte fraktale Brownsche Bewegung verhält. Innerhalb dieses Abschnittes wird nun angenommen, dass der Volatilitätsprozess tatsächlich einer skalierten fraktalen Brownschen Bewegung entspricht und es wird untersucht, wie sich innerhalb dieses Modells Optionen bewerten lassen. Dazu wird zunächst das gesamte Marktmodell unter dem physischen Wahrscheinlichkeitsmaß  $\mathbb{P}$  eingeführt. Anschließend wird ein Maßwechsel zu einem risikoneutralen Maß  $\mathbb{Q}$  durchgeführt. Eine Schwierigkeit, die die fraktale Brownsche Bewegung mit sich bringt, ist die Pfadabhängigkeit, was bedeutet, dass die Verteilung der fraktalen Brownschen Bewegung zu jedem Zeitpunkt vom gesamten vergangenen Pfad anhängig ist. Dies macht es daher im Allgemeinen schwierig den Volatilitätsprozess zu simulieren. Bei dem Wechsel zum äquivalenten Martingalmaß wird sich jedoch herausstellen, dass alle notwendigen Informationen über die Vergangenheit bereits in einem sogenannten Volatilitätsswap eingepreist sind. Es wird weiterhin bewiesen werden, dass der Preisprozess unter dem äquivalenten Martingalmaß  $\mathbb{Q}$  tatsächlich ein Martingal und das Modell somit arbitragefrei ist. In einem weiteren Unterabschnitt wird noch dargestellt, wie sich das Marktmodell unter  $\mathbb{Q}$  simulieren lässt, um beispielsweise Optionen mit Hilfe von Monte Carlo Simulationen zu bewerten. Zuletzt wird noch darauf eingegangen, wie sich das vorgestellte Modell anhand von Marktdaten kalibrieren lässt. Die Schwierigkeit hierbei ist, dass die Simulation des Modells äußerst rechenintensiv ist und nicht in angemessener Zeit durchgeführt werden kann. Es wird daher ein alternativer Ansatz vorgestellt, welcher die Kalibrierung des Modells mit Hilfe von künstlichen neuronalen Netzen deutlich beschleunigen kann. Die Ausführungen in den ersten beiden Unterabschnitten folgen dabei im Groben der Darstellung in [19]. Aus Gründen der Einfachheit wird für den risikolosen Zinssatz  $r$  angenommen, dass  $r = 0$  gilt.

### 5.1 Die Volatilität unter dem physischen Maß $\mathbb{P}$

Im Folgenden wird angenommen, dass sich die Log-Volatilität wie eine skalierte fraktale Brownsche Bewegung verhält. Sei also  $\sigma_t$  der Volatilitätsprozess eines Assets und  $(B_t^H)$  eine fraktale Brownsche Bewegung. Dann wird angenommen, dass für alle  $s, t > 0$

$$\log(\sigma_t) - \log(\sigma_s) = \nu(B_t^H - B_s^H)$$

gilt, wobei  $\nu > 0$ . Im Folgenden wird nun die in Abschnitt 2.1 vorgestellte Mandelbrot-Van-Ness-Darstellung verwendet, um die fraktale Brownsche Bewegung zu definieren. Sei dazu  $(B^\mathbb{P})$  eine zweiseitige Brownsche Bewegung auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ . Dann setze

$$B_t^H = C_H \int_{-\infty}^t (t-u)^{H-1/2} - (-u)_+^{H-1/2} dB_u^\mathbb{P}$$

und die Filtration  $(\mathcal{F}_t)_{t \geq 0}$  sei gegeben durch  $\mathcal{F}_t = \sigma((B_s^\mathbb{P})_{s \in (-\infty, t]}) \vee \sigma(\mathcal{N})$ . Hierbei gilt für  $\sigma$ -Algebren  $\mathcal{A}, \mathcal{B}$  und eine Familie von  $\sigma$ -Algebren  $(\mathcal{A}_i)_{i \in I}$

$$\begin{aligned} \mathcal{A} \vee \mathcal{B} &:= \sigma(\mathcal{A} \cup \mathcal{B}) \\ \bigvee_{i \in I} \mathcal{A}_i &:= \sigma\left(\bigcup_{i \in I} \mathcal{A}_i\right) \end{aligned}$$

und  $\mathcal{N}$  ist die Menge aller  $\mathbb{P}$ -Nullmengen in  $\mathcal{F}_\infty := \bigvee_{t \in [0, \infty)} \mathcal{F}_t$ . Sei nun  $v_t := \sigma_t^2$  der Varianzprozesses. Dann gilt für  $t \geq s \geq 0$

$$\begin{aligned} &\log(v_t) - \log(v_s) \\ &= 2\nu(B_t^H - B_s^H) \\ &= 2\nu C_H \left( \int_{-\infty}^t (t-u)^{H-1/2} - (-u)_+^{H-1/2} dB_u^\mathbb{P} - \int_{-\infty}^s (s-u)^{H-1/2} - (-u)_+^{H-1/2} dB_u^\mathbb{P} \right) \\ &= 2\nu C_H \left( \int_{-\infty}^s (t-u)^{H-1/2} - (s-u)^{H-1/2} dB_u^\mathbb{P} + \int_s^t (t-u)^{H-1/2} dB_u^\mathbb{P} \right). \end{aligned}$$

Setze nun für alle  $t \geq s \geq 0$

$$\begin{aligned} Z_s^\mathbb{P}(t) &:= \int_{-\infty}^s (t-u)^{H-1/2} - (s-u)^{H-1/2} dB_u^\mathbb{P} \\ M_s^\mathbb{P}(t) &:= \int_s^t (t-u)^{H-1/2} dB_u^\mathbb{P}. \end{aligned}$$

Dann sind nach Theorem 2.1 die beiden Prozesse  $(Z_s^\mathbb{P}(t))_{t \geq 0}$  und  $(M_s^\mathbb{P}(t))_{t \geq 0}$  für alle  $t \geq s \geq 0$  zentrierte Gaußprozesse. Weiterhin ist für alle  $t \geq s \geq 0$  die Zufallsvariable  $Z_s^\mathbb{P}(t)$  messbar bezüglich  $\mathcal{F}_s$  und  $M_s^\mathbb{P}(t)$  unabhängig von  $\mathcal{F}_s$ . Insbesondere ist die

Varianz von  $M_s^{\mathbb{P}}(t)$  gegeben durch

$$\mathbb{E}^{\mathbb{P}} \left[ (M_s^{\mathbb{P}}(t))^2 \right] = \int_s^t (t-u)^{2H-1} du = \frac{1}{2H} (t-s)^{2H}.$$

Nun setze

$$\tilde{W}_s^{\mathbb{P}}(t) := \sqrt{2H} M_s^{\mathbb{P}}(t)$$

und  $\eta := (2\nu C_H)/\sqrt{2H}$ . Der Prozess  $(\tilde{W}_s^{\mathbb{P}}(t))_{t \geq 0}$  ist eine skalierte Version von  $(M_s^{\mathbb{P}}(t))$  mit Varianz  $(t-s)^{2H}$ . Damit gilt nun

$$\log(v_t) - \log(v_s) = 2\nu C_H Z_s^{\mathbb{P}}(t) + \eta \tilde{W}_s^{\mathbb{P}}(t)$$

und somit insbesondere

$$v_t = v_s e^{2\nu C_H Z_s^{\mathbb{P}}(t) + \eta \tilde{W}_s^{\mathbb{P}}(t)}.$$

Wegen der Unabhängigkeit von  $\tilde{W}_s^{\mathbb{P}}(t)$ , der  $\mathcal{F}_s$ -Messbarkeit von  $v_s$  und  $Z_s^{\mathbb{P}}(t)$  sowie der Tatsache, dass  $\tilde{W}_s^{\mathbb{P}}(t)$  normalverteilt ist, folgt nun

$$\begin{aligned} \mathbb{E}^{\mathbb{P}} [v_t | \mathcal{F}_s] &= \mathbb{E}^{\mathbb{P}} \left[ v_s e^{2\nu C_H Z_s^{\mathbb{P}}(t) + \eta \tilde{W}_s^{\mathbb{P}}(t)} | \mathcal{F}_s \right] \\ &= v_s e^{2\nu C_H Z_s^{\mathbb{P}}(t)} \mathbb{E}^{\mathbb{P}} \left[ e^{\eta \tilde{W}_s^{\mathbb{P}}(t)} \right] \\ &= v_s e^{2\nu C_H Z_s^{\mathbb{P}}(t)} e^{\frac{1}{2} \mathbb{E}^{\mathbb{P}} \left[ (\eta \tilde{W}_s^{\mathbb{P}}(t))^2 \right]} \\ &= v_s e^{2\nu C_H Z_s^{\mathbb{P}}(t) + \frac{1}{2} \mathbb{E}^{\mathbb{P}} \left[ (\eta \tilde{W}_s^{\mathbb{P}}(t))^2 \right]}. \end{aligned}$$

Diese Gleichung lässt sich nun nutzen, um  $v_t$  wie folgt umzuschreiben:

$$\begin{aligned} v_t &= v_s e^{2\nu C_H Z_s^{\mathbb{P}}(t) + \frac{1}{2} \mathbb{E}^{\mathbb{P}} \left[ (\eta \tilde{W}_s^{\mathbb{P}}(t))^2 \right] + \eta \tilde{W}_s^{\mathbb{P}}(t) - \frac{1}{2} \mathbb{E}^{\mathbb{P}} \left[ (\eta \tilde{W}_s^{\mathbb{P}}(t))^2 \right]} \\ &= \mathbb{E}^{\mathbb{P}} [v_t | \mathcal{F}_s] \mathcal{E} \left( \eta \tilde{W}_s^{\mathbb{P}}(t) \right). \end{aligned} \tag{5}$$

Hierbei ist  $\mathcal{E} \left( \eta \tilde{W}_s^{\mathbb{P}}(t) \right) = e^{\eta \tilde{W}_s^{\mathbb{P}}(t) - \frac{1}{2} \eta^2 \mathbb{E}^{\mathbb{P}} \left[ (\tilde{W}_s^{\mathbb{P}}(t))^2 \right]}$  das stochastische Exponential von  $\eta \tilde{W}_s^{\mathbb{P}}(t)$ . Die Darstellung von  $v_t$  in Gleichung (5) ist deshalb interessant, da zu sehen ist, dass die Verteilung von  $v_t$  lediglich von  $\mathbb{E}^{\mathbb{P}} [v_t | \mathcal{F}_s]$  und nicht von der gesamten Vergangenheit  $\mathcal{F}_s$  abhängt, da  $\tilde{W}_s^{\mathbb{P}}(t)$  unabhängig von  $\mathcal{F}_s$  ist. Das ist insofern von Vorteil, dass wenn die Verteilung von  $\mathbb{E}^{\mathbb{P}} [v_t | \mathcal{F}_s]$  bekannt ist, zur Simulation von  $v_t$  die Filtration  $\mathcal{F}_s$  selbst nicht bekannt sein muss, da der Prozess  $\tilde{W}_s^{\mathbb{P}}(t)$  unabhängig

von  $\mathcal{F}_s$  ist. Wenn also  $v_t$  simuliert werden soll, so stecken alle benötigten Informationen über die Vergangenheit  $\mathcal{F}_s$  bereits in  $\mathbb{E}^\mathbb{P}[v_t|\mathcal{F}_s]$  und es ist nicht notwendig den tatsächlichen Pfad von  $(B_u^\mathbb{P})_{u \in (-\infty, s]}$  zu kennen. Wie sich im folgenden Unterabschnitt zeigen wird, ist die bedingte Erwartung von  $v_t$  unter dem äquivalenten Martingalmaß am Kapitalmarkt beobachtbar. Um das Modell unter dem äquivalenten Martingalmaß zu simulieren genügt es also diese Marktinformationen zu verwenden, um das Problem der Pfadabhängigkeit von  $(v_t)$  praktisch zu lösen, da alle nötigen Informationen über die Pfade von  $(v_t)$  bereits in diesen Marktinformationen stecken.

## 5.2 Die Volatilität unter dem äquivalenten Martingalmaß $\mathbb{Q}$

In diesem Unterabschnitt wird nun untersucht werden, was geschieht, wenn vom physischen Maß  $\mathbb{P}$  zu einem äquivalenten Martingalmaß  $\mathbb{Q}$  gewechselt wird. Dies ist notwendig, um Wertpapiere, wie beispielsweise Optionen, bepreisen zu können. Im Allgemeinen kann es hierbei mehrere äquivalente Martingalmäße geben. Aus Gründen der Einfachheit wird angenommen, dass es sich um einen deterministischen Maßwechsel der Gesalt

$$dB_t^\mathbb{P} = dB_t^\mathbb{Q} + \lambda(t)dt$$

handelt mit  $\lambda : [0, \infty) \rightarrow [0, \infty)$ . Dann gilt also für den Varianzprozess  $v_t$

$$\begin{aligned} v_t &= v_s \exp \left( 2\nu C_H Z_s^\mathbb{P}(t) + \eta \tilde{W}_s^\mathbb{P}(t) \right) \\ &= v_s \exp \left( \eta \tilde{W}_s^\mathbb{Q}(t) + 2\nu C_H \left( Z_s^\mathbb{Q}(t) + \int_s^t (t-u)^{H-1/2} \lambda(u) du \right. \right. \\ &\quad \left. \left. + \int_{-\infty}^s ((t-u)^{H-1/2} - (s-u)^{H-1/2}) \lambda(u) du \right) \right), \end{aligned}$$

wobei  $(Z_t^\mathbb{Q})$  und  $(\tilde{W}_t^\mathbb{Q})$  genau so definiert sind wie  $(Z_t^\mathbb{P})$  und  $(\tilde{W}_t^\mathbb{P})$  nur mit  $B^\mathbb{Q}$  als Integrator. Damit folgt nun für den bedingten Erwartungswert von  $v_t$  unter  $\mathbb{Q}$

$$\begin{aligned} &\mathbb{E}^\mathbb{Q}[v_t|\mathcal{F}_s] \\ &= v_s \exp \left( \frac{1}{2} \eta^2 \mathbb{E}^\mathbb{Q} \left[ \left( \tilde{W}_s^\mathbb{Q}(t) \right)^2 \right] + 2\nu C_H \left( Z_s^\mathbb{Q}(t) + \int_s^t (t-u)^{H-1/2} \lambda(u) du \right. \right. \\ &\quad \left. \left. + \int_{-\infty}^s ((t-u)^{H-1/2} - (s-u)^{H-1/2}) \lambda(u) du \right) \right) \end{aligned}$$

$$\begin{aligned}
&= v_s \exp \left( \frac{1}{2} \eta^2 \mathbb{E}^{\mathbb{P}} \left[ \left( \tilde{W}_s^{\mathbb{P}}(t) \right)^2 \right] + 2\nu C_H Z_s^{\mathbb{P}}(t) \right) \exp \left( 2\nu C_H \int_s^t (t-u)^{H-1/2} \lambda(u) du \right) \\
&= \mathbb{E}^{\mathbb{P}}[v_t | \mathcal{F}_s] \exp \left( 2\nu C_H \int_s^t (t-u)^{H-1/2} \lambda(u) du \right)
\end{aligned}$$

und somit gilt auch unter  $\mathbb{Q}$  für den Varianzprozess

$$\begin{aligned}
v_t &= \mathbb{E}^{\mathbb{P}}[v_t | \mathcal{F}_s] \mathcal{E} \left( \eta \tilde{W}_s^{\mathbb{P}}(t) \right) \\
&= \mathbb{E}^{\mathbb{P}}[v_t | \mathcal{F}_s] \exp \left( \eta \tilde{W}_s^{\mathbb{P}}(t) - \mathbb{E}^{\mathbb{P}} \left[ \left( \tilde{W}_s^{\mathbb{P}}(t) \right)^2 \right] \right) \\
&= \mathbb{E}^{\mathbb{P}}[v_t | \mathcal{F}_s] \exp \left( \eta \sqrt{2H} \int_s^t (t-u)^{H-1/2} \lambda(u) du \right) \\
&\quad \cdot \exp \left( \eta \tilde{W}_s^{\mathbb{Q}}(t) - \mathbb{E}^{\mathbb{Q}} \left[ \left( \tilde{W}_s^{\mathbb{Q}}(t) \right)^2 \right] \right) \\
&= \mathbb{E}^{\mathbb{Q}}[v_t | \mathcal{F}_s] \mathcal{E} \left( \eta \tilde{W}_s^{\mathbb{Q}}(t) \right).
\end{aligned} \tag{6}$$

Im Folgenden wird nun ab sofort angenommen, dass  $s = 0$  gilt, was dem aktuellen Beobachtungszeitpunkt entspricht. Im Allgemeinen wäre es wegen der Pfadabhängigkeit von  $v_t$  nun zur Simulation des Modells notwendig, den gesamten vergangenen Pfad von  $(B_t^{\mathbb{Q}})$  zu berücksichtigen. Dieser wäre jedoch praktisch nicht beobachtbar. Mit Hilfe der Darstellung in Gleichung (6) lässt sich dieses Problem jedoch umgehen. Grund hierfür ist, dass der Prozess  $\tilde{W}_0^{\mathbb{Q}}(t)$  unabhängig von  $\mathcal{F}_0$  ist. Das heißt für die Simulation von  $\tilde{W}_0^{\mathbb{Q}}(t)$  sind keine weiteren Informationen über die Vergangenheit notwendig. Der Ausdruck  $\mathbb{E}^{\mathbb{Q}}[v_t | \mathcal{F}_0]$  entspricht gerade dem aktuellen arbitragenfreien Preis eines Wertpapiers, das im Zeitpunkt  $t$  einen Payoff von  $v_t$  besitzt und ist damit deterministisch. Es wird daher die Funktion  $\xi : [0, \infty) \rightarrow [0, \infty)$  definiert als  $\xi(t) = \mathbb{E}[v_t | \mathcal{F}_0]$ . Diese Funktion kann beispielsweise über die Swapraten eines Varianzswaps geschätzt werden. Bei einem Varianzswap vereinbaren zwei Vertragspartner die Zahlung der Differenz aus der realisierten Varianz und einer vertraglich festgelegten Swaprate zu einem festgelegten Zeitpunkt. Da der Marktwert eines solchen Swaps bei Vertragsabschluss stets Null sein muss, entspricht die Swaprate gerade der vom Markt erwarteten realisierten Varianz. Für genauere Ausführungen sei auf [20, chapter 11] verwiesen.

### 5.3 Zur Arbitragefreiheit des Marktmodells

Im vorangegangenen Unterabschnitt wurde gezeigt, wie sich die Varianz eines Preisprozesses unter einem äquivalenten Martingalmaß  $\mathbb{Q}$  verhält und wie sich das Problem der Pfadabhängigkeit mit Hilfe von Varianzsswaps lösen lässt. Um Optionen auf ein Wertpapier bewerten zu können genügt es natürlich nicht, lediglich den Varianzprozess zu modellieren, sondern auch der Preisprozess selbst muss eingeführt werden. Es ergibt sich damit das folgende Modell

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t dZ_t \\ v_t &= \xi_0(t) \mathcal{E} \left( \eta \tilde{W}_t \right) \end{aligned}$$

wobei

$$\tilde{W}_t = \sqrt{2H} \int_0^t (t-u)^{H-1/2} dB_u.$$

Hierbei sind  $(Z_t)_{t \geq 0}$  und  $(B_t)_{t \geq 0}$  zwei Standard Brownsche Bewegungen. Es lässt sich anhand von Marktdaten beobachten, dass in Phasen von steigender (resp. fallender) Volatilität besonders häufig fallende (resp. steigende) Preise auftreten. Volatilitätsänderungen und Preisänderungen sind also negativ miteinander korreliert.[21] Aus diesem Grund wird angenommen, dass die beiden Brownschen Bewegungen miteinander korreliert sind und zwar mit einem Korrelationskoeffizienten von  $\rho \leq 0$ . Wie sich zwei miteinander korrelierte Brownsche Bewegungen erzeugen lassen, zeigt das folgende Theorem.

**Theorem 5.1.** *Sei  $B_t = (B_t^1, B_t^2)$  eine zweidimensionale Brownsche Bewegung auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathbb{F}, \mathbb{P})$  und sei  $(\mathcal{F}_t)_{t \geq 0}$  die von  $(B_t)$  erzeugte erweiterte Brownsche Filtration. Dann ist für  $\rho \in (-1, 1)$  der Prozess*

$$Z_t = \rho B_t^1 + \sqrt{1 - \rho^2} B_t^2 \quad t \in [0, \infty)$$

*ebenfalls eine Brownsche Bewegung und es gilt  $\mathbb{E}[B_t^1 Z_t] = \rho t$ .*

*Beweis.* Zunächst gilt es zu zeigen, dass der Prozess  $(Z_t)$  ebenfalls eine Brownsche Bewegung ist. Nach Levy's Charakterisierung ([7, Theorem 25.28]) ist  $(Z_t)$  eine Brownsche Bewegung, falls sowohl  $(Z_t)$  selbst als auch  $((Z_t)^2 - t)$  ein  $\mathcal{F}_t$ -Martingal ist. Als Linearkombination von zwei Martingalen ist  $(Z_t)$  ebenfalls ein Martingal. Es bleibt also zu zeigen, dass  $((Z_t)^2 - t)$  ebenfalls ein Martingal ist. Adaptiertheit und

Integrierbarkeit folgen hierbei direkt aus der Definition von  $Z_t$ . Es gilt für  $0 \leq s \leq t$

$$\begin{aligned}
& \mathbb{E}[(Z_t)^2 - t | \mathcal{F}_s] \\
&= \mathbb{E}\left[\left(\rho B_t^1 + \sqrt{1 - \rho^2} B_t^2\right)^2 - t \mid \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\rho^2 (B_t^1)^2 + 2\rho\sqrt{1 - \rho^2} B_t^1 B_t^2 + (1 - \rho^2) (B_t^2)^2 - t \mid \mathcal{F}_s\right] \\
&= \rho^2 \mathbb{E}\left[(B_t^1)^2 - t \mid \mathcal{F}_s\right] + 2\rho\sqrt{1 - \rho^2} \mathbb{E}[B_t^1 B_t^2 | \mathcal{F}_s] + (1 - \rho^2) \mathbb{E}\left[(B_t^2)^2 - t \mid \mathcal{F}_s\right] \\
&= \rho^2 \left((B_s^1)^2 - s\right) + 2\rho\sqrt{1 - \rho^2} B_s^1 B_s^2 + (1 - \rho^2) \left((B_s^2)^2 - s\right) \\
&= \left(\rho B_s^1 + \sqrt{1 - \rho^2} B_s^2\right)^2 - s \\
&= (Z_s)^2 - s
\end{aligned}$$

und somit ist  $(Z_t)^2 - t$  ein Martingal. Es bleibt noch zu zeigen, dass  $\mathbb{E}[B_t^1 Z_t] = \rho t$ . Es gilt wegen der Unabhängigkeit von  $(B_t^1)$  und  $(B_t^2)$ :

$$\begin{aligned}
& \mathbb{E}[B_t^1 Z_t] \\
&= \mathbb{E}\left[B_t^1 \left(\rho B_t^1 + \sqrt{1 - \rho^2} B_t^2\right)\right] \\
&= \rho \mathbb{E}\left[(B_t^1)^2\right] + \sqrt{1 - \rho^2} \mathbb{E}[B_t^1 B_t^2] \\
&= \rho t.
\end{aligned}$$

□

Damit sich innerhalb des Modells Optionen adäquat bepreisen lassen, muss der Preisprozess ein  $(S_t)$  Martingal sein. Um diese zeigen zu können, werden zunächst die folgenden zwei Lemmata benötigt.

**Lemma 5.1.** *Sei  $(Y_t)_{t \geq 0}$  ein reellwertiger stochastischer Prozess auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  mit stetigen Pfaden. Sei weiterhin eine Folge von Stoppzeiten  $(\tau_n)_{n \in \mathbb{N}}$  gegeben mit*

$$\tau_n := \inf\{t > 0 : Y_t = n\}.$$

*Dann gilt  $\lim_{n \rightarrow \infty} \tau_n = \infty$  fast sicher.*

*Beweis.* Angenommen es gibt ein  $\omega \in \Omega$  und eine reelle Zahl  $\hat{\tau} \in [0, \infty)$  mit  $\tau_n(\omega) \rightarrow \hat{\tau}$  für  $n \rightarrow \infty$ . Wegen der Stetigkeit von  $Y_t$  folgt dann

$$Y_{\hat{\tau}} = Y_{\lim_{n \rightarrow \infty} \tau_n(\omega)} = \lim_{n \rightarrow \infty} Y_{\tau_n(\omega)} = \lim_{n \rightarrow \infty} n = \infty.$$



Dies ist aber ein Widerspruch dazu, dass  $(Y_t)$  reellwertig ist.  $\square$

**Lemma 5.2.** *Sei  $(B_t)_{t \geq 0}$  eine Brownsche Bewegung auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  und  $(\mathcal{F}_t)_{t \geq 0}$  die erweiterte Brownsche Filtration. Sei weiterhin  $\sigma : \Omega \times [0, \infty) \rightarrow \mathbb{R}$  ein progressiv messbarer stochastischer Prozess mit  $\sigma_t \leq C, \forall t \in [0, T]$  fast sicher. Dann besitzt die SDE*

$$dS_t = \sigma_t S_t dB_t \quad (7)$$

mit Anfangsbedingung  $S_0 = x \in (0, \infty)$  eine eindeutige Lösung. Insbesondere ist diese Lösung ein Martingal.

*Beweis.* Die SDE (7) lässt sich mit  $\tilde{\sigma}(t, x, \omega) := \sigma(t, \omega)x$  schreiben als

$$dS_t = \tilde{\sigma}(t, S_t) dB_t.$$

Nun ist  $\tilde{\sigma}(t, x, \omega)$  jedoch fast sicher lipschitz-stetig, denn es gilt für fast alle  $\omega \in \Omega$  und alle  $x, y \in \mathbb{R}$  sowie  $t \in [0, T]$

$$|\tilde{\sigma}(t, x, \omega) - \tilde{\sigma}(t, y, \omega)| = |\sigma(t, \omega)x - \sigma(t, \omega)y| \leq C|x - y|$$

womit die SDE höchstens eine Lösung besitzt. Eine Lösung lässt sich aber direkt angeben als

$$S_t = S_0 \exp \left( \int_0^t \sigma_s dB_s - \frac{1}{2} \int_0^t \sigma_s^2 ds \right).$$

Zeige nun, dass für  $\sigma_t$  die Novikov Bedingung erfüllt ist, denn daraus folgt, dass dieser Prozess  $S_t$  ein Martingal ist. Es gilt

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T \sigma_s^2 ds \right) \right] \leq \exp \left( \frac{1}{2} TC \right) < \infty$$

und somit ist die Novikov-Bedingung erfüllt und  $(S_t)$  ist ein Martingal.  $\square$

Nun kann gezeigt werden, dass der Preisprozess  $(S_t)$  tatsächlich ein Martingal ist. Das folgende Theorem und der Beweis folgen dabei der Darstellung aus [22, Theorem 1.1].

**Theorem 5.2.** *Sei ein stochastisches Volatilitätsmodell der folgenden Gestalt gege-*

ben:

$$dS_t = \sigma(t, Y_t) S_t dB_t$$

$$Y_t = \int_0^t K(t, s) dZ_s$$

wobei  $Z_t = \rho B_t + \bar{\rho} \bar{B}_t$  mit  $\rho^2 + \bar{\rho}^2 = 1$  und  $(B_t, \bar{B}_t)$  eine 2-dimensionalen Brownsche Bewegung auf einem filtrierten Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  ist. Die Filtration  $(\mathcal{F}_t)$  sei die von  $(B_t, \bar{B}_t)$  erzeugte erweiterte Brownsche Filtration. Sei weiterhin  $K : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$  dergestalt, dass  $Y_t$  ein Gaußprozess mit stetigen Pfaden und  $\sigma : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}_+$  für jedes  $T > 0, a > 0$  stetig und beschränkt ist auf  $[0, T] \times (-\infty, a]$ . Falls zusätzlich  $\rho \leq 0$  gilt, so ist der Prozess  $(S_t)$  ein echtes Martingal.

*Beweis.* Zunächst ist  $(S_t)$  als stochastisches Integral bezüglich einer Brownschen Bewegung ein lokales Martingal. Da  $(S_t)$  nicht negativ ist, ist  $(S_t)$  somit ein Supermartingal. Ein Supermartingal ist genau dann ein Martingal auf  $[0, T]$ , wenn  $\mathbb{E}[S_T] = S_0$ . Sei nun  $\tau_n := \inf\{t > 0 : Y_t = n\}$  eine Stoppzeit. Da  $\sigma$  beschränkt ist auf  $[0, T] \times (-\infty, n]$  ist der Prozess  $(S_{t \wedge \tau_n})_{t \geq 0}$  nach Lemma 5.2 ein Martingal und es gilt

$$S_0 = \mathbb{E}[S_{T \wedge \tau_n}] = \mathbb{E}[S_T \mathbb{1}_{\{T < \tau_n\}}] + \mathbb{E}[S_T \mathbb{1}_{\{\tau_n \leq T\}}].$$

Nach Lemma 5.1 gilt  $\lim_{n \rightarrow \infty} \tau_n = \infty$  fast sicher und daher folgt mit dem Satz von der monotonen Konvergenz:

$$\lim_{n \rightarrow \infty} \mathbb{E}[S_T \mathbb{1}_{\{T < \tau_n\}}] = \mathbb{E}[S_T].$$

Somit gilt

$$S_0 - \mathbb{E}[S_T] = \lim_{n \rightarrow \infty} \mathbb{E}[S_{\tau_n} \mathbb{1}_{\{\tau_n \leq T\}}].$$

Nun wird ein Maßwechsel zu einem Maß  $\hat{\mathbb{P}}^{(n)}$  durchgeführt. Hierbei gilt

$$\frac{d\hat{\mathbb{P}}^{(n)}}{d\mathbb{P}} = \frac{S_{T \wedge \tau_n}}{S_0}.$$

Da  $(S_{t \wedge \tau_n})_{t \geq 0}$  ein Martingal ist, folgt mit Girsanow's Theorem, dass der Prozess

$$\hat{B}_t^{(n)} := B_t - \int_0^{t \wedge \tau_n} \sigma(s, Y_s) ds$$

eine Brownsche Bewegung unter  $\hat{\mathbb{P}}^{(n)}$  ist und es gilt

$$\mathbb{E}[S_{\tau_n} \mathbb{1}_{\{\tau_n \leq T\}}] = \mathbb{E}^{\hat{\mathbb{P}}^{(n)}}[S_0 \mathbb{1}_{\{\tau_n \leq T\}}] = S_0 \hat{\mathbb{P}}^{(n)}(\tau_n \leq T).$$

Nun gilt für  $t \leq \tau_n$

$$\begin{aligned} Y_t &= \int_0^t K(t, s) dZ_s \\ &= \int_0^t \rho K(t, s) dB_s + \int_0^t \bar{\rho} K(t, s) d\bar{B}_s \\ &= \int_0^t \rho K(t, s) \hat{B}_s^{(n)} + \int_0^t \bar{\rho} K(t, s) d\bar{B}_s + \int_0^t \rho K(t, s) \sigma(s, Y_s) ds \\ &= \int_0^t K(t, s) d\bar{Z}_s^{(n)} + \int_0^t \rho K(t, s) \sigma(s, Y_s) ds \\ &=: \hat{Y}_t + \int_0^t \rho K(t, s) \sigma(s, Y_s) ds \end{aligned}$$

wobei  $\bar{Z}_s^{(n)}$  eine Brownsche Bewegung unter  $\bar{\mathbb{P}}^{(n)}$  ist. Da nun  $Y_t \leq \hat{Y}_t$  für  $t \leq \tau_n$ , gilt  $\tau_n \geq \hat{\tau}_n := \inf\{t > 0 : \hat{Y}_t = n\}$  und somit

$$\lim_{n \rightarrow \infty} \hat{\mathbb{P}}^{(n)}(\tau_n \leq T) \leq \lim_{n \rightarrow \infty} \hat{\mathbb{P}}^{(n)}(\hat{\tau}_n \leq T) = \lim_{n \rightarrow \infty} \hat{\mathbb{P}}^{(n)}\left(\sup_{t \in [0, T]} \hat{Y}_t \geq n\right) = 0$$

woraus folgt, dass

$$S_0 - \mathbb{E}[S_T] = 0.$$

Somit ist der Prozess  $(S_t)$  ein echtes Martingal. □

Es bleibt zu zeigen, dass das vorgestellte Marktmodell tatsächlich alle Bedingungen von Theorem 5.3 erfüllt. Die Funktionen  $\sigma$  und  $K$  sind hierbei gegeben durch

$$\begin{aligned} K(t, s) &= \sqrt{2H}(t-s)^{H-\frac{1}{2}} \\ \sigma(t, y) &= \sqrt{\xi_0(t) e^{\eta y - \frac{1}{2}\eta^2 t^{2H}}}. \end{aligned}$$

Die Funktion  $\sigma$  ist genau dann stetig und beschränkt auf  $[0, T] \times (-\infty, a]$  für alle  $T, a > 0$ , falls die Funktion  $\xi_0(t)$  stetig ist und beschränkt auf  $[0, T]$ . Die Funkti-

on  $\xi_0(t)$  wird mit Hilfe von Marktdaten geschätzt und sollte deshalb so modelliert werden, dass sie stetig ist. Es bleibt zu zeigen, dass der Prozess  $\tilde{W}_t = \int_0^t K(t, s)dB_s$  eine Version mit stetigen Pfaden besitzt. Dazu wird das folgende Theorem aus [23] verwendet.

**Theorem 5.3.** *Sei  $Y_t$  ein Gaußprozess der Gestalt*

$$Y_t = \int_0^t K(t, s)dB_s$$

*wobei  $(B_t)$  eine Brownsche Bewegung ist und  $K(t, \cdot) \in L^2([0, t])$  für alle  $t \geq 0$ . Dann besitzt der Prozess  $(Y_t)$  eine Version mit  $\gamma$ -hölderstetigen Pfaden für alle  $\gamma < \alpha$ , falls Konstanten  $\alpha, C > 0$  existieren, so dass die folgenden drei Bedingungen erfüllt sind:*

1. *Die Funktion  $K(t, s)$  ist stetig differenzierbar in  $t$  für alle  $0 < s < t$*
2.  *$|\frac{\partial K}{\partial t}(t, s)| \leq C(t - s)^{\alpha - \frac{3}{2}}$  für alle  $0 < s < t$*
3.  *$\int_s^t K(t, u)^2 du \leq C(t - s)^{2\alpha}$  für alle  $0 < s < t$*

*Beweis.* Siehe [23, S. 779] □

Nun gilt für die Funktion  $K(t, s) = \sqrt{2H}(t - s)^{H - \frac{1}{2}}$  aber

$$\left| \frac{\partial K}{\partial t}(t, s) \right| = \sqrt{2H} \left| H - \frac{1}{2} \right| (t - s)^{H - \frac{3}{2}}$$

sowie

$$\int_s^t K(t, u)^2 du = (t - s)^{2H}.$$

Damit sind die Bedingungen von Theorem 5.3 für  $C = 1$  und  $\alpha = H$  erfüllt und der Prozess  $(\tilde{W}_t)$  besitzt eine Version mit  $\gamma$ -hölderstetigen Pfaden für alle  $\gamma < H$ .

Das Marktmodell ist damit nach Theorem 5.3 tatsächlich arbitragefrei und somit kann der Wert einer Option bestimmt werden, indem der erwartete Payoff der Option unter dem äquivalenten Martingalmaß bestimmt wird. Dazu bietet es sich an, diesen Erwartungswert mit Hilfe von Monte-Carlo-Simulationen zu berechnen. Dafür ist es notwendig, sowohl den Preisprozess als auch den Volatilitätsprozess zu simulieren. Im folgenden Unterabschnitt wird gezeigt, wie eine solche Simulation des Modells durchgeführt werden kann.

## 5.4 Simulation des Modells

Nachdem im vorangegangenen Abschnitt gezeigt wurde, wie das Marktmodell unter dem äquivalenten Martingalmaß  $\mathbb{Q}$  aussieht, stellt sich nun die Frage, wie sowohl die Pfade der Varianz als auch des Preisprozesses simuliert werden können. Das Marktmodell ist gegeben durch die folgende Dynamik:

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t dZ_t \\ v_t &= \xi(t) \mathcal{E} \left( \eta \tilde{W}_t \right) \\ \tilde{W}_t &= \sqrt{2H} \int_0^t (t-u)^{H-\frac{1}{2}} dB_u. \end{aligned}$$

Hierbei sind  $(Z_t)$  und  $(B_t)$  zwei Brownsche Bewegungen, die negativ korreliert sind mit einem Korrelationskoeffizienten  $\rho \leq 0$ . Die Funktion  $\xi : \mathbb{R} \rightarrow \mathbb{R}$  ist eine stetige Funktion, die mit Hilfe von Swapraten geschätzt werden kann.

Zunächst gilt es die Pfade der Varianz  $\tilde{W}_t$  und  $(Z_t)$  zu simulieren. Der Prozess  $(\tilde{W}_t)$  ist als ein stochastisches Integral mit einem deterministischen, quadratintegrierbarem Integranden ein Gaußprozess. Auch  $(Z_t)$  ist als Brownsche Bewegung insbesondere ein Gaußprozess. Um die Pfade der beiden Prozesse simulieren zu können, genügt es daher die gemeinsame Kovarianzmatrix zu bestimmen. Dafür gilt es die Kovarianzfunktion von  $\tilde{W}_t$  sowie die Korrelation zwischen den beiden Prozessen zu berechnen. Die Kovarianzfunktion der Brownschen Bewegung ist gegeben durch  $\mathbb{E}[Z_t Z_s] = \min\{t, s\}$ .

**Theorem 5.4.** *Die Kovarianzfunktion von  $(\tilde{W}_t)$  ist für  $t \geq s$  gegeben durch*

$$\mathbb{E}[\tilde{W}_t \tilde{W}_s] = s^{2H} G\left(\frac{s}{t}\right)$$

mit

$$G(x) = \frac{2H}{H + \frac{1}{2}} x^{\frac{1}{2}-H} {}_2F_1\left(\frac{1}{2} - H, 1, \frac{3}{2} + H, x\right),$$

wobei

$${}_2F_1(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 u^{b-1} (1-u)^{c-b-1} (1-zu)^{-a} du$$

die Gaußsche hypergeometrische Funktion ist.

*Beweis.* Mit der zweiten binomischen Formel folgt zunächst

$$\mathbb{E}[\tilde{W}_t \tilde{W}_s] = \frac{1}{2} \left( \mathbb{E}[(\tilde{W}_t)^2] + \mathbb{E}[(\tilde{W}_s)^2] - \mathbb{E}[(\tilde{W}_t - \tilde{W}_s)^2] \right),$$

und es gilt

$$\begin{aligned} \mathbb{E}[(\tilde{W}_t)^2] &= 2H \cdot \mathbb{E} \left[ \left( \int_0^t (t-s)^{H-\frac{1}{2}} dB_s \right)^2 \right] \\ &= 2H \int_0^t (t-s)^{2H-1} ds \\ &= t^{2H}. \end{aligned}$$

Weiterhin gilt

$$\begin{aligned} &\mathbb{E}[(\tilde{W}_t - \tilde{W}_s)^2] \\ &= 2H \mathbb{E} \left[ \left( \int_0^t (t-u)^{H-\frac{1}{2}} dB_s - \int_0^s (s-u)^{H-\frac{1}{2}} dB_u \right)^2 \right] \\ &= 2H \left( \int_0^s \left( (t-u)^{H-\frac{1}{2}} - (s-u)^{H-\frac{1}{2}} \right)^2 du + \int_s^t (t-u)^{2H-1} du \right) \\ &= 2H \int_0^s (t-u)^{2H-1} + (s-u)^{2H-1} - 2(t-u)^{H-\frac{1}{2}}(s-u)^{H-\frac{1}{2}} du + (t-s)^{2H} \\ &= -(t-s)^{2H} + t^{2H} + s^{2H} - 4H \int_0^s (t-u)^{H-\frac{1}{2}}(s-u)^{H-\frac{1}{2}} du + (t-s)^{2H} \\ &= t^{2H} + s^{2H} - 4H \int_0^s (t-u)^{H-\frac{1}{2}}(s-u)^{H-\frac{1}{2}} du \end{aligned}$$

und daraus folgt nun

$$\begin{aligned} \mathbb{E}[\tilde{W}_t \tilde{W}_s] &= 2H \int_0^s (t-u)^{H-\frac{1}{2}}(s-u)^{H-\frac{1}{2}} du \\ &= 2H t^{H-\frac{1}{2}} s^{H-\frac{1}{2}} \int_0^s \left(1 - \frac{u}{t}\right)^{H-\frac{1}{2}} \left(1 - \frac{u}{s}\right)^{H-\frac{1}{2}} du \\ &= 2H t^{H-\frac{1}{2}} s^{H+\frac{1}{2}} \int_0^1 \left(1 - \frac{s}{t}z\right)^{H-\frac{1}{2}} (1-z)^{H-\frac{1}{2}} dz \\ &= s^{2H} \left(\frac{s}{t}\right)^{\frac{1}{2}-H} \frac{2H}{H+\frac{1}{2}} {}_2F_1\left(\frac{1}{2}-H, 1, \frac{3}{2}+H, \frac{s}{t}\right) \\ &= s^{2H} G\left(\frac{s}{t}\right), \end{aligned}$$

wobei  $z = \frac{u}{s}$  substituiert und  $\Gamma(x+1) = x\Gamma(x)$  verwendet wurde. □

Um auch den Preisprozess simulieren zu können, muss nun auch ein Pfad der Brownschen Bewegung  $(Z_t)$  erzeugt werden. Hierfür ist es notwendig, die Korrelation zwischen  $(Z_t)$  und  $(\tilde{W}_t)$  zu bestimmen. Dazu wird angenommen, dass  $Z_t = \rho B_t + \bar{\rho} \bar{B}_t$  wobei  $\bar{\rho} = \sqrt{1 - \rho^2}$  und  $(\bar{B}_t)$  eine Brownsche Bewegung unabhängig von  $(B_t)$  ist.

**Theorem 5.5.** *Für alle  $t \geq s$  gilt*

$$\mathbb{E}[\tilde{W}_t Z_s] = \rho \frac{\sqrt{2H}}{H + \frac{1}{2}} \left( t^{H+\frac{1}{2}} - (t-s)^{H+\frac{1}{2}} \right)$$

sowie

$$\mathbb{E}[\tilde{W}_s Z_t] = \rho \frac{\sqrt{2H}}{H + \frac{1}{2}} s^{H+\frac{1}{2}}.$$

*Beweis.* Mit der ersten binomischen Formel folgt

$$\mathbb{E}[\tilde{W}_t Z_s] = \frac{1}{2} \left( (\tilde{W}_t + Z_s)^2 - \mathbb{E}(\tilde{W}_t)^2 - \mathbb{E}(Z_s)^2 \right).$$

Nun gilt

$$\begin{aligned} & \mathbb{E}(\tilde{W}_t + Z_s)^2 \\ &= \mathbb{E} \left[ \left( \sqrt{2H} \int_0^t (t-u)^{H-\frac{1}{2}} dB_u + \rho B_s + \bar{\rho} \bar{B}_s \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \sqrt{2H} \int_0^t (t-u)^{H-\frac{1}{2}} dB_u + \int_0^s \rho dB_u \right)^2 \right] + (\bar{\rho})^2 s \\ &= \left[ \left( \sqrt{2H} \int_0^s (t-u)^{H-\frac{1}{2}} + \frac{\rho}{\sqrt{2H}} dB_u + \sqrt{2H} \int_s^t (t-u)^{H-\frac{1}{2}} dB_u \right)^2 \right] + (1-\rho^2)s \\ &= 2H \int_0^s \left( (t-u)^{H-\frac{1}{2}} + \frac{\rho}{\sqrt{2H}} \right)^2 du + 2H \int_s^t (t-u)^{2H-1} du + (1-\rho^2)s \\ &= 2H \int_0^s (t-u)^{2H-1} + 2\rho \frac{(t-u)^{H-\frac{1}{2}}}{\sqrt{2H}} + \frac{\rho^2}{2H} du + 2H \int_s^t (t-u)^{2H-1} du + (1-\rho^2)s \\ &= 2H \int_0^t (t-u)^{2H-1} du + 2\rho\sqrt{2H} \int_0^s (t-u)^{H-\frac{1}{2}} du + \rho^2 s + (1-\rho^2)s \\ &= t^{2H} + 2\rho \frac{\sqrt{2H}}{H + \frac{1}{2}} \left( t^{H+\frac{1}{2}} - (t-s)^{H+\frac{1}{2}} \right) + s \end{aligned}$$

und da  $\mathbb{E}[(\tilde{W}_t)^2] = t^{2H}$  und  $\mathbb{E}[(Z_s)^2] = s$ , folgt

$$\mathbb{E}[\tilde{W}_t Z_s] = \rho \frac{\sqrt{2H}}{H + \frac{1}{2}} \left( t^{H+\frac{1}{2}} - (t-s)^{H+\frac{1}{2}} \right).$$

Nun gilt außerdem

$$\begin{aligned}
& \mathbb{E} \left[ (\tilde{W}_s + Z_t)^2 \right] \\
&= \mathbb{E} \left[ \left( \sqrt{2H} \int_0^s (s-u)^{H-\frac{1}{2}} dB_u + \rho B_t + \bar{\rho} \bar{B}_t \right)^2 \right] \\
&= \mathbb{E} \left[ \left( \sqrt{2H} \int_0^s (s-u)^{H-\frac{1}{2}} + \frac{\rho}{\sqrt{2H}} dB_u + \int_s^t \rho dB_u \right)^2 + (\bar{\rho})^2 t \right] \\
&= 2H \int_0^s \left( (s-u)^{H-\frac{1}{2}} + \frac{\rho}{\sqrt{2H}} \right)^2 du + \rho^2(t-s) + (1-\rho^2)t \\
&= 2H \int_0^s (s-u)^{2H-1} + \frac{\rho^2}{2H} + 2\frac{\rho}{\sqrt{2H}}(s-u)^{H-\frac{1}{2}} du - \rho^2 s + t \\
&= s^{2H} + \rho^2 s + 2\rho \frac{\sqrt{2H}}{H + \frac{1}{2}} s^{H+\frac{1}{2}} - \rho^2 s + t \\
&= s^{2H} + 2\rho \frac{\sqrt{2H}}{H + \frac{1}{2}} s^{H+\frac{1}{2}} + t
\end{aligned}$$

und damit folgt

$$\mathbb{E}[\tilde{W}_s Z_t] = \rho \frac{\sqrt{2H}}{H + \frac{1}{2}} s^{H+\frac{1}{2}}.$$

□

Damit können nun Pfade von  $\tilde{W}_t$  und  $Z_t$  simuliert werden. Sei dazu  $T > 0$  der Zeitpunkt, bis zu dem simuliert werden soll und  $N \in \mathbb{N}$  die Anzahl der Zeitpunkte zu den simuliert wird und  $h = \frac{T}{N}$  die Schrittweite. Nun können die Pfade in den folgenden drei Schritten simuliert werden.

### Schritt 1:

Zunächst wird die Kovarianzmatrix  $\Sigma^{\tilde{W}}$  von  $\tilde{W}_t$  mit  $\Sigma_{i,j}^{\tilde{W}} = \mathbb{E} \left[ \tilde{W}_{ih} \tilde{W}_{jh} \right]$  bestimmt. Ebenso wird die Kovarianzmatrix  $\Sigma^Z$  gebildet. Zuletzt wird die Kovarianzmatrix zwischen  $\tilde{W}$  und  $Z$  berechnet, welche gegeben ist durch  $\Sigma_{i,j}^{\tilde{W},Z} = \mathbb{E} \left[ \tilde{W}_{ih} Z_{jh} \right]$ . Aus diesen Matrizen kann nun die gemeinsame Kovarianz  $\Sigma$  bestimmt werden, welche gegeben ist durch

$$\Sigma = \begin{pmatrix} \Sigma^{\tilde{W}} & \Sigma^{\tilde{W},Z} \\ (\Sigma^{\tilde{W},Z})^T & \Sigma^Z \end{pmatrix}.$$

### Schritt 2:

Nun wird ein Vektor  $v \in \mathbb{R}^{2N}$  von unabhängigen und standardnormalverteilten



Zufallsvariablen generiert. Daraus ergeben sich dann die Pfade von  $\tilde{W}$  und  $Z$  als

$$\begin{pmatrix} (\tilde{W}_{ih})_{i=1,\dots,N} \\ (Z_{jh})_{j=1,\dots,N} \end{pmatrix} = \Sigma \cdot v.$$

### Schritt 3:

Im dritten und letzten Schritt werden nun die Pfade des Preisprozesses  $S_t$  simuliert. Dazu kann das Euler-Schema verwendet werden. Sei dazu  $S_0 > 0$  gegeben. Dann setze für  $i = 1, \dots, N$ :

$$\begin{aligned} S_{ih} &= S_{(i-1)h} + \sqrt{\sigma_{(i-1)h}}(Z_{ih} - Z_{(i-1)h}) \\ &= S_{(i-1)h} + \sqrt{\xi((i-1)h)} \exp\left(\frac{1}{2}\eta\tilde{W}_{(i-1)h} - \frac{1}{2}\eta^2((i-1)h)^{2H}\right)(Z_{ih} - Z_{(i-1)h}) \end{aligned}$$

## 5.5 Kalibrierung des Modells

Es wurde nun also ein Marktmodell vorgestellt, welches das empirische Verhalten des Volatilitätsprozesses sehr gut beschreiben kann. Weiterhin wurde gezeigt, dass dieses Modell tatsächlich arbitragefrei ist und wie sich der Varianz- sowie der Preisprozess simulieren lassen. Damit können sich nun theoretisch beliebige Optionen innerhalb dieses Modells bewerten lassen. Das Ganze kann mit Hilfe von Monte-Carlo-Simulationen getan werden. Hierbei wird eine große Anzahl an Pfaden des Preisprozesses simuliert, für jeden Pfad der Payoff der Option bestimmt, und anschließend der Mittelwert aus diesen Payoffs berechnet. Dabei stellt sich nun jedoch die Frage, wie die drei Parameter  $H, \eta$  und  $\rho$  dieses Modells gewählt werden sollen. Eine Möglichkeit wäre es, diese Daten auf Basis von historischen Daten zu schätzen, wie es für die Parameter  $H$  und  $\eta$  im Abschnitt 4 getan wurde. Das Problem hierbei ist jedoch, dass diese Schätzungen nicht unbedingt mit den vom Markt angenommenen Werten übereinstimmen müssen, da die Marktteilnehmer auch zukünftige Entwicklungen bereits antizipieren.

### 5.5.1 Anpassung an die Volatilitätsfläche

Eine andere Möglichkeit die Parameter des Modells zu bestimmen, ist die Anpassung des Modells an die aktuellen Marktdaten. Es wird bei diesem Vorgang auch von der Modellkalibrierung gesprochen. Hierbei werden die Parameter so gewählt, dass die vom Modell generierten fairen Preise von Optionen mit denen am Markt möglichst gut übereinstimmen. Typischerweise werden hierfür europäische Optionen mit verschiedenen Basispreisen und Laufzeiten verwendet. Anstatt das Modell so zu kalibrieren, dass die vom Modell erzeugten Preise mit denen am Markt beobachte-

ten bestmöglich übereinstimmen, können genau so gut die impliziten Volatilitäten des Modells mit denen der für die Kalibrierung verwendeten Optionen in möglichst gute Übereinstimmung gebracht werden. An dieser Stelle ist die sogenannte Volatilitätsoberfläche, welche in Abschnitt 3.2 eingeführt wurde, erneut von Bedeutung. Ziel der Kalibrierung des Modells ist es nämlich nun, die Parameter des Modells so zu bestimmen, dass die Volatilitätsoberfläche des Modells möglichst gut mit der vom Markt generierten Volatilitätsoberfläche übereinstimmt. Möglichst gute Übereinstimmung bedeutet hierbei typischerweise, dass der durchschnittliche quadratische Abstand zwischen der vom Modell generierten und der vom Markt gegebenen Volatilitätsoberfläche minimal ist. Der Vorteil eines stochastischen Volatilitätsmodells, im Gegensatz zum Black-Scholes-Modell mit konstanter Volatilität, ist die Tatsache, dass dieses ebenfalls eine Volatilitätsoberfläche erzeugt, welche nicht konstant ist und verschiedene Parameter zu unterschiedlichen Volatilitätsoberflächen führen.

Wie Bayer, Friz und Gatheral in [19, Kapitel 5] zeigten, lässt sich bereits mit einer einfachen Schätzung der Parameter eine Volatilitätsoberfläche erzeugen, die sehr gut zu den Marktdaten passt. Dass die Parameter in diesem Artikel lediglich geschätzt wurden, hat den einfachen Grund, dass die Simulation des Modells äußerst rechenintensiv ist. Bereits um die implizite Volatilität für eine Option zuverlässig zu bestimmen, müssen mehrere Tausend Pfade des Preisprozesses simuliert werden. Bereits dies ist sehr rechenintensiv und nimmt viel Zeit in Anspruch. Nun muss jedoch die Volatilitätsoberfläche, das heißt die implizite Volatilität für mehrere Optionen, bestimmt werden. Da weiterhin eine bestmögliche Parameterkombination gesucht ist, muss diese Volatilitätsoberfläche auch noch für unterschiedlichste Modellparameter berechnet werden. Hier wird klar, was der große Nachteil eines stochastischen Volatilitätsmodells wie des rBergomi-Modells ist: es ist kaum möglich das Modell mit einer begrenzten Rechenleistung in angemessener Zeit zu kalibrieren. Einen Ausweg aus diesem Problem bieten künstliche neuronale Netze, welche in der Lage sind, den Zusammenhang zwischen Modellparametern und Volatilitätsoberfläche für ein gegebenes stochastisches Volatilitätsmodell zu lernen.

### **5.5.2 Kalibrierung mit Hilfe künstlicher neuronaler Netze**

In diesem Unterabschnitt soll kurz erläutert werden, wie die Kalibrierung des rBergomi-Modells, oder auch jedes anderen Marktmodells, mit Hilfe von künstlichen neuronalen Netzen drastisch beschleunigt werden kann. Hierbei werden nur die Grundideen vorgestellt und es wird nicht darauf eingegangen, wie beispielsweise der Lernprozess

eines neuronalen Netzwerks abläuft. Für eine Einführung in die Theorie künstlicher neuronaler Netze sei beispielsweise auf [24] verwiesen. Die Darstellungen in diesem Unterabschnitt folgen den Darstellungen in [25] und [26].

Sei  $(\mathcal{M}(\theta))_{\theta \in \Theta}$  eine Menge von Marktmodellen, wobei  $\Theta \subseteq \mathbb{R}^m$  mit  $m \in \mathbb{N}$  die Menge aller möglichen Parameter des Modells sei. Weiterhin sei  $Z$  die Menge aller möglichen Optionsparameter. Typischerweise werden bei der Modellkalibrierung lediglich europäische Optionen betrachtet. Die zur Bepreisung relevanten Informationen sind in dem Fall die Restlaufzeit  $T$  und die log-Moneyness  $k$ , also haben die  $\zeta \in Z$  in diesem Fall die Gestalt  $\zeta = (T, k)$ . Falls also ausschließlich europäische Optionen betrachtet werden, so ist  $Z \subseteq \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ . Nun erzeugt ein gegebenes Marktmodell stets eine bestimmte Volatilitätsoberfläche. Die Abbildung, die gegebenen Modell- und Optionsparametern die implizite Volatilität zuordnet sei gegeben durch

$$\begin{aligned} \sigma_{BS}^{\mathcal{M}(\theta)} : Z &\rightarrow \mathbb{R}_{\geq 0} \\ \zeta &\rightarrow \sigma_{BS}^{\mathcal{M}(\theta)}(\zeta). \end{aligned}$$

Sei nun weiterhin  $Z' \subset Z$  die Menge an Optionsparametern für diejenigen Optionen, für die Marktpreise vorliegen und die zur Kalibrierung des Marktmodells verwendet werden sollen. Sei weiterhin  $\sigma_{BS}^{MKT} : Z' \rightarrow \mathbb{R}_{\geq 0}$  diejenige Abbildung, die den Optionsparametern die entsprechende, am Markt beobachtete, implizite Volatilität zuordnet. Das Ziel der Kalibrierung ist es nun diejenigen Modellparametern  $\theta^*$  zu finden, die für eine gegebene Abstandsfunktion  $\delta$  den Abstand der vom Markt gegebenen und den vom Modell generierten impliziten Volatilitäten zu minimieren, das heißt

$$\theta^* = \operatorname{argmin}_{\theta \in \Theta} \delta \left( \left( \sigma_{BS}^{\mathcal{M}(\theta)}(\zeta) \right)_{\zeta \in Z'}, \left( \sigma_{BS}^{MKT}(\zeta) \right)_{\zeta \in Z'} \right). \quad (8)$$

Häufig wird als Abstandsfunktion eine gewichtete quadratische Abweichung verwendet, das heißt

$$\theta^* = \operatorname{argmin}_{\theta \in \Theta} \sum_{\zeta \in Z'} w_{\zeta} \left( \sigma_{BS}^{\mathcal{M}(\theta)}(\zeta) - \sigma_{BS}^{MKT}(\zeta) \right)^2$$

wobei  $w_{\zeta} > 0$  für  $\zeta \in Z'$  bestimmt, wie stark die einzelnen, zur Kalibrierung verwendeten, Optionen gewichtet werden. Hier kann beispielsweise festgelegt werden, dass Optionen mit einem hohen Bid-Ask-Spread eher untergewichtet werden. Typischerweise ist nun die Funktion  $\sigma_{BS}^{\mathcal{M}(\theta)}$  nicht explizit bekannt und muss mit Hilfe von

Monte-Carlo-Simulationen für eine Auswahl von Modellparametern geschätzt werden. Anschließend werden diejenigen Modellparameter ausgewählt, für die der Abstand der Volatilitätsoberflächen minimal ist. Wie bereits erwähnt tritt im rBergomi-Modell nun jedoch das Problem auf, dass eben jene Monte-Carlo-Simulation wegen der Pfadabhängigkeit äußerst rechenintensiv ist. Um diese Problematik zu lösen ist nun die Idee, die Funktion  $\sigma_{BS}^{\mathcal{M}(\theta)}$  durch ein künstliches neuronales Netz zu approximieren. Das neuronale Netze soll also bei einem überwachten Lernvorgang den funktionalen Zusammenhang zwischen Modellparametern und Volatilitätsoberfläche approximieren. Anschließend kann im Minimierungsproblem (8) die Funktion  $\sigma_{BS}^{\mathcal{M}(\theta)}$  durch die Approximation ersetzt werden. Dies hat den Vorteil, dass lediglich im ersten Schritt beim Trainieren des neuronalen Netzes Monte-Carlo-Simulationen durchgeführt werden müssen. Da das neuronale Netz lediglich die Volatilitätsoberfläche des gegebenen Modells für verschiedene Modellparameter approximiert und diese nicht von der aktuellen Marktsituation abhängig ist, muss dieser Schritt nur einmalig durchgeführt werden. Anschließend kann diese Approximation immer wieder zur Kalibrierung des Modells herangezogen werden und es müssen keine neuen Monte-Carlo-Simulationen mehr durchgeführt werden.

Wie der Lernvorgang des neuronalen Netzes abläuft, soll im Folgenden kurz skizziert werden. Zunächst wird ein festes Gitter über die Volatilitätsoberfläche gelegt. Dieses Gitter sei gegeben durch  $\zeta_i = (T_i, k_i)$  für  $i = 1, \dots, N$ , wobei  $N \in \mathbb{N}$ . Nun wird für eine Auswahl an Modellparametern  $\theta \in \hat{\Theta}$ ,  $|\hat{\Theta}| = N_{train} \in \mathbb{N}$  die implizite Volatilität an diesen Stützstellen mit Hilfe von Monte-Carlo-Simulationen berechnet. Damit ergibt sich dann für jeden ausgewählten Parameter  $\theta \in \hat{\Theta}$  die Menge  $\{\sigma_{BS}^{\mathcal{M}(\theta)}(T_i, k_i) : i = 1, \dots, N\}$ . Mit Hilfe dieses generierten Datensatzes kann nun ein Feed-Forward-Netz trainiert werden, dessen Architektur in Abbildung 17 zu sehen ist. Der Input des neuronalen Netzes entspricht also gerade den Modellparametern. Ziel des Trainingsvorgangs ist es, die Gewichte auf den Kanten des neuronalen Netzwerks so anzupassen, dass der Output des Netzes gerade der impliziten Volatilitätsoberfläche an den Stützstellen für den gegebenen Input entspricht. Hierfür kann beispielsweise der Back-Propagation-Algorithmus verwendet werden. Sei  $\tilde{F}(w; \theta)$  der Output des neuronalen Netzes mit Input  $\theta$  und  $w \in \mathbb{R}^M$  seien die Gewichte auf den Verbindungen zwischen den Neuronen, wobei  $M \in \mathbb{N}$  die Gesamtanzahl an Kanten sei. Da es sich um ein Feed-Forward-Netz handelt, existieren lediglich Kanten von einer Schicht in die Nächste.

Nun wird also in einem ersten Schritt das neuronale Netz trainiert, was bedeutet,

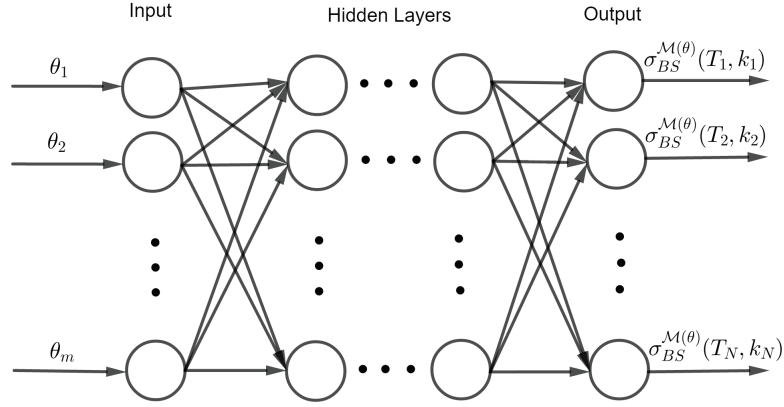


Abbildung 17: Architektur des neuronalen Netzes

dass das folgende Optimierungsproblem gelöst werden muss:

$$w^* = \operatorname{argmin}_{w \in \mathbb{R}^M} \sum_{\theta \in \Theta} \sum_{i=1}^N \left( \tilde{F}(w; \theta) - \sigma_{BS}^{\mathcal{M}(\theta)}(T_i, k_i) \right)^2.$$

Um diese Zielfunktion zu minimieren wird typischerweise der Backpropagation-Algorithmus verwendet. Für eine genauere Erläuterung dieses Algorithmus sei auf [27] verwiesen. Nachdem dieser Schritt abgeschlossen wurde, wird als nächstes das Optimierungsproblem (8) wie folgt umgeschrieben:

$$\theta^* = \operatorname{argmin}_{\theta \in \Theta} \delta \left( \tilde{F}(w^*; \theta), (\sigma_{BS}^{MKT}(\zeta))_{\zeta \in Z'} \right).$$

Es wurde also die tatsächliche Volatilitätsfläche des Modells ersetzt durch die Approximation des neuronalen Netzes. Auch hier bietet es sich an, wieder einen gewichteten quadratischen Abstand zwischen den Volatilitätsflächen als Abstandsfunktion zu verwenden. Zur numerischen Lösung dieses Optimierungsproblems schlagen Bayer et al den sogenannten Levenberg-Marquardt-Algorithmus vor [26, S. 8]. Für eine ausführliche Beschreibung dieses Algorithmus sei auf [28] und [29] verwiesen.

Es stellt sich noch die Frage, wie die Modellparameter  $\theta \in \Theta$  innerhalb des rBergomi-Modells definiert werden. Offensichtlich sind die Parameter des Modells der Hurst-Parameter  $H$ , der Korrelationskoeffizient  $\rho$ , der Parameter  $\eta$  und die Funktion  $\xi(t)$ . Die ersten drei Parameter sind offensichtlich reelle Zahlen und lassen sich damit direkt als Input für das neuronale Netz verwenden. Auch die Funktion  $\xi(t)$  kann, wenn sie nicht über Varianzswaps geschätzt wird, als Input für das neuronale Netz

verwendet werden, indem sie als konstant angenommen oder als Treppenfunktion modelliert wird.

## 6 Fazit

Ziel dieser Arbeit ist es gewesen, die von Gatheral, Jaisson und Rosenbaum präsentierten Eigenschaften des Volatilitätsprozesses anhand von zwei ausgewählten Datensätzen zu verifizieren und das sich daraus ergebende Marktmodell im Hinblick auf Arbitragefreiheit und Kalibrierung zu untersuchen. Das zentrale mathematische Objekt, welches zur Modellierung des Volatilitätsprozesses in dieser Arbeit verwendet wurde, ist die fraktale Brownsche Bewegung. Dieser Prozess wurde in einem ersten Abschnitt formal definiert und es wurde gezeigt, dass sich dieser als ein stochastisches Integral bezüglich einer zweiseitigen Standard Brownschen Bewegung darstellen lässt. Es wurden außerdem charakteristische Eigenschaften wie die Selbstähnlichkeit, die Stationarität der Inkremente und die Hölder-Stetigkeit der Pfade besprochen. Da sich lediglich der Preis, nicht aber die Volatilität eines Assets, am Markt beobachten lässt, wurden im zweiten Abschnitt Konzepte zur Schätzung des Volatilitätsprozesses präsentiert. Es stellte sich heraus, dass die implizite Volatilität nicht dazu geeignet ist, da diese von den Parametern der betrachteten Option abhängt. Dies bestätigte die These, dass die Volatilität eines Assets nicht als konstant angenommen werden kann. Als alternatives Konzept zur Schätzung des Volatilitätsprozesses wurde daher die realisierte Volatilität eingeführt, welche sich mit Hilfe von hochfrequenten Marktdaten berechnen lässt.

Nachdem eine Methode zur Schätzung des Volatilitätsprozesses vorgestellt wurde, wurde dieser für zwei verschiedene Assets, einen Aktienindex und eine Einzelaktie, genauer untersucht. Es stellte sich dabei heraus, dass die logarithmierte realisierte Volatilität beider Assets dieselbe Skalierungseigenschaft wie eine skalierte fraktale Brownsche Bewegung mit einem Hurstparameter von etwa 0.1 aufweist. Aufgrund des niedrigen Hurst-Parameters, welcher ein Maß für die Regularität der Pfade ist, wird die Volatilität auch als raue Volatilität bezeichnet. Eine genauere Analyse der Inkremente der Log-Volatilität deutete jedoch drauf hin, dass diese nicht normalverteilt sind, wie es bei einer fraktalen Brownschen Bewegung der Fall wäre, und insbesondere extreme Anstiege der Volatilität deutlich häufiger zu beobachten sind. Nichtsdestotrotz wurde anschließend angenommen, dass die Log-Volatilität eines Assets gerade einer skalierten fraktalen Brownschen Bewegung entspricht. Es wurde ein deterministischer Maßwechsel vom physischen zu einem äquivalenten Martingalmaß durchgeführt und es stellt sich heraus, dass die Verteilung des zukünftigen Volatilitätsprozesses lediglich von der vom Markt antizipierten zukünftigen Varianz abhängt. Diese lässt sich jedoch mit Hilfe der Swapraten von sogenannten Varianzswaps schätzen. Anschließend wurde gezeigt, dass der Preisprozess in diesem

Modell tatsächlich ein Martingal und der Markt damit arbitragefrei ist.

Zuletzt wurde gezeigt, wie sich das eingeführte Modell simulieren und mit Hilfe der vom Markt gegebenen Volatilitäts Oberfläche kalibrieren lässt. Da die Simulation des Modells äußerst rechen- und eine Kalibrierung auf klassischem Wege dadurch zu zeitintensiv ist, wurde noch ein alternativer Ansatz zur Kalibrierung des Modells mit Hilfe von künstlichen neuronalen Netzen präsentiert.



## Literaturverzeichnis

- [1] F. Black and M. Scholes, “The pricing of options and corporate liabilities,” *Journal of political economy*, vol. 81, no. 3, pp. 637–654, 1973.
- [2] S. L. Heston, “A closed-form solution for options with stochastic volatility with applications to bond and currency options,” *The review of financial studies*, vol. 6, no. 2, pp. 327–343, 1993.
- [3] J. Gatheral, T. Jaisson, and M. Rosenbaum, “Volatility is rough,” *Quantitative Finance*, vol. 18, no. 6, pp. 933–949, 2018.
- [4] B. B. Mandelbrot and J. W. Van Ness, “Fractional brownian motions, fractional noises and applications,” *SIAM review*, vol. 10, no. 4, pp. 422–437, 1968.
- [5] J. Picard, “Representation formulae for the fractional brownian motion,” in *Séminaire de Probabilités XLIII*, pp. 3–70, Springer, 2011.
- [6] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang, *Stochastic calculus for fractional Brownian motion and applications*. Springer Science & Business Media, 2008.
- [7] A. Klenke, *Probability theory: a comprehensive course*. Springer Science & Business Media, 2nd ed., 2013.
- [8] G. Shevchenko, “Fractional brownian motion in a nutshell,” *arXiv preprint arXiv:1406.1956*, 2014.
- [9] R. B. Davies and D. Harte, “Tests for hurst effect,” *Biometrika*, vol. 74, no. 1, pp. 95–101, 1987.
- [10] J. E. Gentle, *Computational statistics*, vol. 308. Springer, 2009.
- [11] B. Mandelbrot, “The variation of certain speculative prices,” *The Journal of Business*, vol. 36, no. 4, pp. 394–419, 1963.
- [12] F. Black and M. Scholes, “The valuation of option contracts and a test of market efficiency,” *The journal of finance*, vol. 27, no. 2, pp. 399–417, 1972.
- [13] M. McAleer and M. C. Medeiros, “Realized volatility: A review,” *Econometric Reviews*, vol. 27, no. 1-3, pp. 10–45, 2008.
- [14] M. Capiński, E. Kopp, and J. Traple, *Stochastic calculus for finance*. Cambridge University Press, 2012.

- [15] O. E. Barndorff-Nielsen and N. Shephard, “Estimating quadratic variation using realized variance,” *Journal of Applied econometrics*, vol. 17, no. 5, pp. 457–477, 2002.
- [16] F. M. Bandi and J. R. Russell, “Microstructure noise, realized variance, and optimal sampling,” *The Review of Economic Studies*, vol. 75, no. 2, pp. 339–369, 2008.
- [17] L. von Auer, *Ökonometrie*. Springer Science & Business Media, 3rd ed., 2005.
- [18] C. M. Jarque and A. K. Bera, “A test for normality of observations and regression residuals,” *International Statistical Review/Revue Internationale de Statistique*, pp. 163–172, 1987.
- [19] C. Bayer, P. Friz, and J. Gatheral, “Pricing under rough volatility,” *Quantitative Finance*, vol. 16, no. 6, pp. 887–904, 2016.
- [20] J. Gatheral, *The volatility surface: a practitioner’s guide*, vol. 357. John Wiley & Sons, 2011.
- [21] J. Bae, C.-J. Kim, and C. R. Nelson, “Why are stock returns and volatility negatively correlated?,” *Journal of Empirical Finance*, vol. 14, no. 1, pp. 41–58, 2007.
- [22] P. Gassiat, “On the martingale property in the rough bergomi model,” *Electronic Communications in Probability*, vol. 24, no. 33, pp. 1–9, 2019.
- [23] E. Alos, O. Mazet, D. Nualart, *et al.*, “Stochastic calculus with respect to gaussian processes,” *The Annals of Probability*, vol. 29, no. 2, pp. 766–801, 2001.
- [24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016.
- [25] B. Horvath, A. Muguruza, and M. Tomas, “Deep learning volatility,” *Available at SSRN: <https://ssrn.com/abstract=3322085>*, 2019.
- [26] C. Bayer, B. Horvath, A. Muguruza, B. Stemper, and M. Tomas, “On deep calibration of (rough) stochastic volatility models,” *arXiv preprint arXiv:1908.08806*, 2019.
- [27] R. Hecht-Nielsen, “Theory of the backpropagation neural network,” in *Neural networks for perception*, pp. 65–93, Elsevier, 1992.

- [28] K. Levenberg, “A method for the solution of certain non-linear problems in least squares,” *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164–168, 1944.
- [29] D. W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

## Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Seitens des Verfassers bestehen keine Einwände die vorliegende Masterarbeit für die öffentliche Benutzung im Universitätsarchiv zur Verfügung zu stellen.

Jena, 30.09.2020